ZAMAN SERİSİ REGRESYONLARINDA ARDIŞIK BAĞINTI ve DEĞİŞEN VARYANS

Hüseyin Taştan[1]

[1]Yıldız Teknik Üniversitesi
İktisat Bölümü

Ders Kitabı:
Introductory Econometrics: A Modern Approach (4th ed.)
J. Wooldridge

25 Nisan 2013

- Bölüm 11'de gördüğümüz gibi, herhangi bir modelin dinamik özellikleri doğru belirlendiği takdirde hata teriminde ardışık bağıntının (serial correlation) olmaması gerekir.
- Bununla beraber bazı statik modellerde ve FDLM modellerinde spesifikasyon hatası olmasa bile hata terimlerinde otokorelasyon olabileceğini gördük.
- Bu bölümde "ardışık bağıntı"nın (serial correlation) ne tür problemlere yol açacağını ve ne gibi önlemler alabileceğimizi öğreneceğiz.

Bölüm Planı

- Hata teriminde ardışık bağıntı varken OLS tahmin edicilerin özellikleri
- Ardışık bağıntı testleri
- x değişkenleri kesin dışsal iken ardışık bağıntı probleminin giderilmesi
- Farkı alınmış seriler kullanıldığında ardışık bağıntının düzeltilmesi
- Ardışık bağıntı varken OLS standart hataları ve test istatistiklerinin düzeltilmesi
- Zaman serilerinde değişen varyans (ARCH, GARCH modelleri)

Hata terimlerinde ardışık bağıntı (serial correlation) varken OLS tahmin edicilerinin özellikleri nasıldır?

- CH 10'da TS.1-TS.3 varsayımları altında OLS'nin sapmasız olduğunu gördük. Burada sapmasızlığı sağlayan x'lerin kesin dışsal (strictly exogenous) oluşuydu. Hata teriminde otokorelasyon olması önemli değildi.
- Değişen varyans OLS t.e.nin sapmasızlığına zarar vermiyordu. Benzer şekilde kesin dışsallık altında serial correlation da sapmasızlığı zedelemez.
- CH 11'de kesin dışsallık varsayımının yerine "verinin zayıf bağımlı (weakly dependent) olması" koşulunu koyarak yumuşattık. Bu halde OLS sapmasız olmasa bile tutarlı **(consistent)** oluyordu ve bu sonuç ardışık bağımlılık'tan etkilenmiyordu.

Etkinlik (efficiency) ve hipotez testleri

- Teorem 10.4 (Gauss-Markov teoremi) BLUE için hem homoscedasticity hem de "hatalarda ardışık bağıntı olmaması" koşulunu getiriyordu. O halde, hata teriminde otokorelasyon olması durumunda OLS tahmin edicileri artık BLUE değildir. Test istatistikleri ve se'ler asimptotik olarak bile geçerli değildir.
- Bunu, Gauss-Markov varsayımlarının ilk 4'ü geçerli ve artıklar AR(1) süreci izliyor iken OLS tahmin edicilerin varyansını hesaplayarak gösterebiliriz.

$$u_t = \rho u_{t-1} + e_t, \quad t = 1, 2, ..., n \tag{1}$$

$$|\rho| < 1 \tag{2}$$

Etkinlik (efficiency) ve hipotez testleri

- e_t, sıfır ortalamalı ve sabit varyanslı ilişkisiz bir rassal değişkendir (uncorrelated random variable). (2), daha önce gördüğümüz gibi kararlılık (stability) koşuludur.
- Basit bir regresyonda β_1 eğim katsayısının OLS tahmin edicisini şöyle yazabiliriz :

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- Formülü basitleştirmek adına, x_t'nin örneklem ortalamasının sıfır olduğunu varsayalım, $\bar{x} = 0$. Buna göre β_1'in OLS tahmin edicisi $\hat{\beta}_1$ şöyle yazılabilir:

$$\hat{\beta}_1 = \beta_1 + SST_x^{-1} \sum_{t=1}^{n} x_t u_t \tag{3}$$

- Burada $SST_x = \sum_{t=1}^{n} x_t^2$. Şimdi, $\hat{\beta}_1$'in (X'e koşullu) varyansını hesaplarken, u_t'deki ardışık bağımlılığı göz önünde bulundurmalıyız.

Etkinlik (efficiency) ve hipotez testleri

$$
\begin{aligned}
\mathsf{Var}(\hat{\beta}_1) &= SST_x^{-2}\mathsf{Var}\left(\sum_{t=1}^{n} x_t u_t\right) \\
&= SST_x^{-2}\left(\sum_{t=1}^{n} x_t^2 \mathsf{Var}(u_t) + 2\sum_{t=1}^{n-1}\sum_{j=1}^{n-t} x_t x_{t+j}\mathsf{E}(u_t u_{t+j})\right) \\
&= \frac{\sigma^2}{SST_x} + 2\left(\frac{\sigma^2}{SST_x^2}\right)\sum_{t=1}^{n-1}\sum_{j=1}^{n-t} \rho^j x_t x_{t+j} \qquad (4)
\end{aligned}
$$

- Denklemde, $\sigma^2 = \mathsf{Var}(u_t)$'dir. Ayrıca $\mathsf{E}(u_t u_{t+j}) = \mathsf{Cov}(u_t, u_{t+j}) = \rho^j \sigma^2$ olduğunu kullandık. (4) nolu denklemdeki ilk terim, σ^2/SST_x, $\rho = 0$ iken $\hat{\beta}_1$'in varyansıdır ki Gauss-Markov varsayımları altında bildiğimiz OLS varyansına eşittir.

Etkinlik (efficiency) ve hipotez testleri

- (4)'de ilk terim Gauss-Markov varsayımları altında $\rho = 0$ durumunda ki standart OLS varyansıdır. Eğer serial correlation varsa ($\rho \neq 0$), (4)'de ikinci terimi ihmal etmemiz varyans tahmininde sapmaya yol açacaktır.
- Pek çok ekonomik seride $\rho > 0$ ve x'ler zaman içinde kendi geçmişleriyle pozitif yönde ilişkili olduğu için (4)'deki ikinci terimin ihmal edilmesi varyansın olduğundan daha düşük tahmin edilmesine (underestimation) yol açacaktır.
- $\hat{\beta}$'lerin se'leri bu varyansların kare köküne eşit oluğu için varyansdaki sapma tüm test istatistiklerini (t, F, LM) geçersiz kılacaktır.
- Düşük hesaplanan varyans, daha yüksek t değerleri demek olduğu için, aslında sıfırdan farklı olmayan katsayıların anlamlı imiş gibi çıkmalarına sebep olacaktır.

- Ekonometri kitapları çoğu kez "regresyonda gecikmeli bağımlı değişken varsa ve hata terimleri otokorelasyon içeriyorsa OLS tutarsızdır" şeklinde ifadeler içerir. Bu çok muğlak bir ifadedir ve doğru değildir. Daha kesin (precise) bir ifade bulmamız gerekir.
- Bunu göstermek için, y_{t-1} veriyken y_t'nin beklenen değerinin doğrusal olduğunu varsayalım:

$$E(y_t|y_{t-1}) = \beta_0 + \beta_1 y_{t-1} \tag{5}$$

- Burada kararlılığı varsayıyoruz, yani $|\beta_1| < 1$. Bu ifadeyi bir hata terimi ile yazabiliriz:

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t \tag{6}$$

$$E(u_t|y_{t-1}) = 0 \tag{7}$$

- Kurulumumuz gereği bu model, OLS'nin tutarlılığı için gerekli olan TS.3' varsayımını sağlar, böylece OLS tahmin edicileri, $\hat{\beta}_0$ ve $\hat{\beta}_1$, tutarlıdır.

- Burada $\{u_t\}$'nin otokorelasyona sahip olmasını yasaklayan bir koşul yoktur.
- Koşul (7) u_t ile y_{t-1}'in ilişkisiz olması gerektiğini söylüyor, ancak u_t ile y_{t-2} ilişkili olabilir.
- $u_{t-1} = y_{t-1} - \beta_0 - \beta_1 y_{t-2}$ olduğu için $Cov(u_t, u_{t-1}) = -\beta_1 Cov(u_t, y_{t-2})$'ye eşittir ve sıfır olması için bir zorunluluk yoktur.
- Yani, hata terimlerinin serial correlation içerdiği ve y_{t-1}'in bağımsız değişken olduğu (6)'daki regresyonda OLS $\hat{\beta}$'ları tutarlıdır (consistent). Çünkü bu β'lar (5)'deki koşullu beklenen değer denkleminin parametreleridir.
- Bu durumda serial correlation varyansın sapmalı çıkmasına ve test istatistiklerinin geçersiz olmasına yol açacak, ancak tutarlılığı **(consistency)** etkilemeyecektir.

- Peki, y_{t-1}'in açıklayıcı değişken olarak yer aldığı ve hata teriminin otokorelasyon içerdiği bir regresyonda $\hat{\beta}$ katsayıları ne zaman tutarsız (inconsistent) olacaktır ?
- (6)'daki regresyonda hata terimi u_t'nin (1)'deki gibi bir AR(1) süreci izlemesi durumunda OLS $\hat{\beta}$'ları tutarsız olacaktır.

$$E(e_t|u_{t-1}, u_{t-2}, ...) = E(e_t|y_{t-1}, y_{t-2}, ...) = 0 \qquad (8)$$

- e_t, varsayım gereği y_{t-1} ile ilişkisiz olduğundan, $Cov(y_{t-1}, u_t) = \rho Cov(y_{t-1}, u_{t-1})$ olur ki bu ifade $\rho = 0$ olmadıkça sıfıra eşit değildir. Bu ise y_t'nin y_{t-1} üzerine regresyonundan gelen β_0 ve β_1'in OLS tahmin edicilerinin tutarsız (inconsistent) olmasına neden olur.

Bağımlı Değişkenin Gecikmesi Varken Ardışık Bağıntı

- Ancak, burada şu nokta gözden kaçmamalı: Eğer (6)'da u_t AR(1) süreci izliyorsa, bu, y_t'nin bir AR(2) süreci izlediğini, dolayısıyla da (6) modelinde spesifikasyon hatası olduğunu gösterir. Bunu (6) ile (1)'i birleştirerek görebiliriz.
- $u_{t-1} = y_{t-1} - \beta_0 - \beta_1 y_{t-2}$ ifadesini $u_t = \rho u_{t-1} + e_t$ içine koyarsak (6) şöyle yazılabilir:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 y_{t-1} + \rho(y_{t-1} - \beta_0 - \beta_1 y_{t-2}) + e_t \\ &= \beta_0(1-\rho) + (\beta_1 + \rho) y_{t-1} - \rho\beta_1 y_{t-2} + e_t \\ &= \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + e_t \end{aligned}$$

- Denklemde $\alpha_0 = \beta_0(1-\rho)$, $\alpha_1 = \beta_1 + \rho$,ve $\alpha_2 = -\rho\beta_1$'dir. (8) veriyken, aşağıdaki ifade yazılabilir:

$$E(y_t|y_{t-1}, y_{t-2}, \ldots) = E(y_t|y_{t-1}, y_{t-2}) = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} \tag{9}$$

- Buna göre, y'nin bütün geçmiş değerleri veriyken, y_t'nin beklenen değeri iki gecikmesine bağlıdır.

- AR(2) modelinin kararlılık (stability) koşulları (Section 12.3'de göreceğiz) altında, (9)'un OLS tahminleri tutarlı (consistent) ve asimptotik olarak normal dağılmış tahmin edicilerdir.
- Demek ki, regresyona hem gecikmeli bağımlı değişken koyup hem de hataların belli bir model izlediğini varsaydığımız durumlarda çok dikkatli olmalıyız. Aksi halde, tutarsız tahmin ediciler elde edebiliriz.
- Dinamik modellerde hata teriminin serial correlation içermesi çoğu kez modelimizin eksik spesifikasyonuna delalet eder.
- Yukarıdaki örnekte (6)'daki regresyona y_{t-2} değişkenini ekleyince sorun kalmadı.

- İlk önce x'lerin kesin dışsal (**strictly exogenous**) olduğu durumda çoklu regresyonda hata terimlerinin **serial correlation** içerip içermediğini teşhise yönelik çeşitli testler göreceğiz.
- Kesin dışsallık, u_t'nin x'lerin tüm dönemlere ait (geçmiş, şimdiki zaman ve gelecek) değerleriyle ilişkisiz olduğu anlamına geliyordu. **Dolayısıyla, gecikmeli bağımlı değişken kullanılan modellere bu testler uygulanamaz.**

u_t, AR(1) izlesin ve x'ler kesin dışsal:

- Hata teriminde çeşitli biçimlerde ardışık bağıntı olabilir. En basit model, hata teriminin (1) ve (2)'deki gibi AR(1) süreci izlemeleri halidir.

- Kesin dışsallık varsayımı dolayısıyla u_t'nin beklenen değeri x'lerin tüm geçmiş, cari ve gelecek değerlerine koşullu olarak sıfırdır :

$$E(e_t|u_{t-1}, u_{t-2}, ...) = 0 \tag{10}$$

$$Var(e_t|u_{t-1}) = Var(e_t) = \sigma_e^2 \tag{11}$$

- Bunlar AR(1) modelindeki standart varsayımlardır ($\{e_t\}$ i.i.d bir seri olduğunda), ve bunlar Bölüm 11'deki büyük örneklem sonuçlarını dinamik regresyona uygulamamıza olanak sağlar.

$$H_0 : \rho = 0. \tag{12}$$

Kesin Dışsal Açıklayıcı Değişkenler Varken AR(1) Ardışık Bağıntı Testi

1. y_t'nin $x_{t1}, ..., x_{tk}$ üzerine regresyonundan $t = 1, 2, ..., n$ için OLS kalıntılarını, $\hat{u}_t$, elde edin.
2. Şu regresyonu tahmin edin

$$\hat{u}_t \text{ 'nin } \quad \hat{u}_{t-1} \quad \text{üzerine,} \quad t = 2, ..., n \tag{13}$$

   ve $\hat{u}_{t-1}$'in katsayısı $\hat{\rho}$'yu ve ilgili t istatistiğini hesaplayın, $t_{\hat{\rho}}$.
3. $H_0 : \rho = 0$ boş hipotezini $H_0 : \rho \neq 0$ alternatifine karşı test etmek için $t_{\hat{\rho}}$'yi kullanın.(Aslında beklenen $\rho > 0$ olduğundan alternatif hipotez $H_0 : \rho > 0$ olabilir.) Karar ilgili p-değerine göre verilir.

Example 12.1: Phillips Eğrisi için AR(1) Ardışık Bağıntı Testi

- Statik Phillips eğrisi için, (13)'daki regresyondan $\hat{\rho} = 0.573$ ve $t = 4.93$, p-değeri$= 0.000$ (48 gözlem) bulunur. Bu ise, 1. dereceden pozitif ardışık bağıntı lehinde güçlü kanıt sunar.
- Bunun bir sonucu, Bölüm 10'daki se ve t istatistiklerinin geçersiz olmasıdır.
- Tersine beklentilerle genişletilmiş Phillips eğrisi için $\hat{\rho} = -0.036$ ve $t = -0.297$, p-değeri$= 0.775$ (47 gözlem) bulunur. Buna göre, beklentilerle genişletilmiş Phillips eğrisinde AR(1) ardışık bağıntı olduğu yönünde kanıt yoktur.

Klasik Varsayımlar altında Durbin-Watson Testi

- AR(1) ardışık bağıntı için bir diğer test Durbin-Watson testidir. **Durbin-Watson (DW) istatistiği** OLS kalıntılarına dayanır:
$$DW = \frac{\sum_{t=2}^{n} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{n} \hat{u}_t^2} \qquad (14)$$
- Cebirsel olarak (13)'deki $\hat{\rho}$ ile DW bağıntılıdır:
$$DW \approx 2(1 - \hat{\rho}) \qquad (15)$$
- Bunun tam eşitlik olmama nedeni $\hat{\rho}$'nun paydasında $\sum_{t=2}^{n} \hat{u}_{t-1}^2$ varken DW istatistiğinin paydasında OLS kalıntılarının kareleri toplamının olmasıdır. Ancak makul örneklem büyüklüklerinde (15)'deki yakınsama geçerlidir. Bu durumda, DW testi ile $\hat{\rho}$ için t testi kavramsal olarak aynıdır.

Durbin-Watson Testi

- Durbin-Watson (1950), CLM varsayımları altında DW istatistiğinin dağılımını türetti.
- Hata terimlerinin normal dağıldığı varsayımı da dahil tüm CLM varsayımlarının sağlanması gerekmektedir.
- Orjinal regresyonda sabit terim (intercept) olmak zorundadır.
- Yine, regresyonda açıklayıcı değişkenler arasında gecikmeli bağımlı değişkenler, $y_{t-1}, y_{t-2}, \ldots$, olmamalıdır.
- Kritik değerler gözlem saysı n ve x sayısına (k) göre değişen alt (dL) ve üst (dU) limitlerden oluşmaktadır.

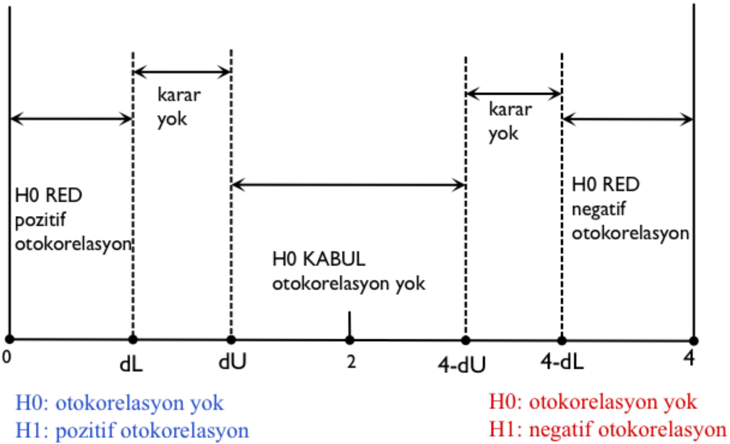
Durbin-Watson Testi

- DW testi, genellikle, aşağıdaki alternatif hipotez için uygulanır:

$$H_1 : \rho > 0 \tag{16}$$

- (15)'deki yakınsama, $\hat{\rho} \approx 0$ iken $DW \approx 2$ ve $\hat{\rho} > 0$ iken $DW < 2$ olduğunu ifade eder. (12)'deki boş hipotezi (16) lehine reddedebilmek için 2'den istatistiksel açıdan anlamlı olarak küçük olan bir DW değeri arıyoruz.
- DW'nin boş hipotez altında dağılımını elde etmedeki sorunlar nedeniyle, DW'yi iki kritik değer ile kıyaslamalıyız. Bunlar, dU (üst limit) ve dL (alt limit) olarak gösterilir. $DW < dL$ ise H_0'ı (16) lehine reddederiz; $DW > dU$ ise H_0'ı reddedemeyiz. $dL \leq DW \leq dU$ ise test kararsızdır.

Durbin-Watson Testi



Kesin Dışsal Açıklayıcı Değişkenler Yokken AR(1) Ardışık Bağıntıyı Test Etme

- Açıklayıcı değişkenler kesin dışsal (**strictly exogenous**) değilse, bu durumda herhangi bir x_{tj}, u_{t-1} ile ilişkili olabileceğinden ne (13)'deki t testi ne de DW testi geçerli olacaktır.
- Hatta büyük örneklerde bile bunlar geçerli olmayacaktır.
- Kesin dışsallığın olmadığı modellere en iyi örnek gecikmeli bağımlı değişkenlerin, y_{t-1}, bulunduğu modellerdir. Bu modelde y_{t-1} ile u_{t-1} açıktır ki ilişkili olacaklardır.
- x'lerden bazılarının kesin dışsal olmadığı durumda Durbin'in şu alternatif testi uygulanır :

1. y_t'nin $x_{t1},...,x_{tk}$ üzerine regresyonundan $t = 1, 2, ..., n$ için OLS kalıntılarını, $\hat{u}$, elde edin.
2. Şu regresyonu tahmin edin

$$\hat{u}_t \text{ 'nin } \quad x_{t1},...,x_{tk},\hat{u}_{t-1} \quad \text{üzerine,} \quad t = 2,...,n \qquad (17)$$

   $\hat{u}_{t-1}$'in katsayısı $\hat{\rho}$'yı ve ilgili t istatistiğini hesaplayın, $t_{\hat{\rho}}$.
3. $H_0 : \rho = 0$ boş hipotezini $H_1 : \rho \neq 0$ alternatifine karşı test etmek için $t_{\hat{\rho}}$'yi kullanın (Veya tek taraflı alternatife karşı test edin).

- (17)'de OLS kalıntılarının tüm açıklayıcı değişkenler, sabit ve kalıntıların gecikmesi üzerine regresyonunu hesapladık. Kalıntıların gecikmesi üzerindeki t istatistiği AR(1) modelinde (12)'nin test edilmesinde geçerlidir (H_0 altında $Var(u_t|x_t, u_{t-1}) = \sigma^2$ olduğunu varsayarsak). Bağımlı değişkenin gecikmeleri ve kesin dışsal olmayan açıklayıcı değişkenler de x_{tj}'ler arasında yer alabilir.

Örnek 12.2: Asgari Ücret Denklemi için AR(1) Ardışık Bağıntı Testi

- Arkaplandaki stokastik süreçlerin zayıf bağımlı olduğunu varsayıyoruz, ancak doğrusal bir zaman trendi içermelerine izin veriyoruz. $\hat{u}_t$, OLS kalıntıları olmak üzere, 37 gözlemle aşağıdaki regresyon hesaplanmıştır.

$\hat{u}_t$ 'nin $log(mincov_t), log(prgnp_t), log(usgnp_t), t, \hat{u}_{t-1}$ üzerine,

- $\hat{u}_{t-1}$'nin katsayı tahmini $\hat{\rho} = 0.481$ ve $t = 2.89$, p-değeri$= 0.007$. Buna göre hatalarda AR(1) ardışık bağıntı yönünde güçlü kanıt bulunmaktadır, yani, $\hat{\beta}_j$'ya ait t istatistikleri geçersizdir, bunlara göre çıkarım yapılamaz. Hatırlarsak, u_t'ler açıklayıcı değişkenlerle cari olarak ilişkisiz ise $\hat{\beta}_j$'lar hala tutarlıdır. Görülüyor ki, bunun yerine (13)'deki regresyonu kullansaydık $\hat{\rho} = 0.417$ ve $t = 2.63$ elde edecektik, yani bu örnek için sonuçlar benzerdir.

Yüksek Dereceden Ardışık Bağıntıları Test Etme

- (17)'deki test kolaylıkla daha yüksek dereceden ardışık bağıntılar için geliştirilebilir. Aşağıdaki AR(2) Modelini ele alalım:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + e_t$$

- Bu model için aşağıdaki boş hipotezi test etmek istiyoruz:

$$H_0 : \rho_1 = 0, \rho_2 = 0 \tag{18}$$

- Bu alternatif model ikinci dereceden ardışık bağıntıyı test etmek için kullanılabilir. Her zaman ki gibi, önce OLS kalıntılarını, $\hat{u}_t$, elde edeceğiz. Sonrasında ise aşağıdaki regresyonu tahmin edeceğiz.

$$\hat{u}_t \text{ 'nin} \quad x_{t1}, ..., x_{tk}, \hat{u}_{t-1}, \hat{u}_{t-2} \quad \text{üzerine,} \quad t = 3, ..., n$$

- Buradan $\hat{u}_{t-1}$ ve $\hat{u}_{t-2}$'nin birlikte anlamlılığı için F istatistiğini elde edebiliriz. Eğer bu iki gecikme, makul anlamlılık düzeylerinde, anlamlılarsa, (18)'i reddedebilir ve hataların ardışık bağımlı olduğuna karar verebiliriz.
- Genelleştirirsek, q. dereceden otoregresif modeldeki ardışık bağımlılığı da test edebiliriz:

$$u_t = \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots + \rho_q u_{t-q} + e_t \tag{19}$$

- Boş hipotez ise:

$$H_0 : \rho_1 = 0, \rho_2 = 0, \ldots, \rho_q = 0 \tag{20}$$

AR(q) Ardışık Bağıntının Test Edilmesi

1. y_t'nin $x_{t1}, ..., x_{tk}$ üzerine regresyonundan $t = 1, 2, ..., n$ için OLS kalıntılarını, $\hat{u}$, elde edin.
2. Şu regresyonu tahmin edin

$$\hat{u}_t' nin \;\; x_{t1}, ..., x_{tk}, \hat{u}_{t-1}, \hat{u}_{t-2}, ..., \hat{u}_{t-q} \;\; uzerine, \;\; t = (q+1), ..., n \tag{21}$$

3. $\hat{u}_{t-1}$, $\hat{u}_{t-2}$,...,$\hat{u}_{t-q}$'nun birlikte anlamlılığını test etmek için F istatistiğini hesaplayın.

- Eğer x_{tj}'ler kesin dışsalsa, yani her x_{tj}, $u_{t-1}, u_{t-2}, ..., u_{t-q}$ ile ilişkisizse, x_{tj}'ler (21) nolu regresyondan çıkarılabilir. Regresyona x_{tj}'leri eklemek kesin dışsallık varsayımı olsun veya olmasın testin geçerliliğini bozmaz. Test sabit varyans varsayımını gerektirmektedir:

$$Var(u_t | x_t, u_{t-1}, ..., u_{t-q}) = \sigma^2 \tag{22}$$

AR(q) Ardışık Bağıntının Test Edilmesi

- F testini hesaplamanın bir alternatifi Lagrange Çarpanı (LM) istatistiğini kullanmaktır. (20)'yi test için kullanılacak LM istatistiği şöyle hesaplanabilir:

$$LM = (n - q)R^2_{\hat{u}} \tag{23}$$

- $R^2_{\hat{u}}$, (21) nolu regresyonun belirlilik katsayısıdır. Boş hipotez altında LM istatistiği χ^2_q dağılımını takip eder. Bu teste AR(q) ardışık bağıntı için **Breusch-Godfrey testi** denilmektedir.
- LM istatistiği de sabit varyans varsayımını gerektirmektedir, ancak değişen varyansa dirençli (robust) hale getirilebilir (Wooldridge (1991b)).

Kesin Dışsal Açıklayıcı Değişkenler Varken Ardışık Bağıntının Düzeltilmesi

- Hata terimi AR(1) süreci izliyor ve değişkenlerimiz kesin dışsal iseler ardışık bağıntı durumunda test istatistiklerinde bazı düzeltmeler yapabiliriz.
- Basitlik amacıyla, tek açıklayıcı değişken olan durumu düşünelim:

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad t = \ 1, 2, ..., n$$

- Denklemdeki sorun, u_t'deki ardışık bağıntı olduğundan, ardışık bağıntıyı ortadan kaldıracak şekilde denklemin dönüştürülmesi gerekmektedir. $t \geq 2$ için :

$$y_{t-1} = \beta_0 + \beta_1 x_{t-1} + u_{t-1}$$

$$y_t = \beta_0 + \beta_1 x_t + u_t$$

- İlk denklemi ρ ile çarpıp ikincisinden çıkarırsak

$$y_t - \rho y_{t-1} = (1-\rho)\beta_0 + \beta_1(x_t - \rho x_{t-1}) + e_t,\ t \geq 2$$

- $e_t = u_t - \rho u_{t-1}$ eşitliğini kullandık. Bu son denklemi aşağıdaki gibi yazabiliriz:

$$\tilde{y}_t = (1-\rho)\beta_0 + \beta_1 \tilde{x}_t + e_t,\ t \geq 2 \qquad (24)$$

- Bu dönüştürmede ise aşağıdakiler fark-benzeri seri adını alır (**quasi-differenced data**):

$$\tilde{y}_t = y_t - \rho y_{t-1}, \tilde{x}_t = x_t - \rho x_{t-1} \qquad (25)$$

- Eğer $\rho = 1$ ise bunlar farkı alınmış serilerdir ancak $\rho < 1$ olduğunu varsayıyoruz. (24)'deki hata terimleri ardışık olarak ilişkisizdir, gerçekten de, bu denklem bütün Gauss-Markov varsayımlarını sağlamaktadır.

Kesin Dışsal Açıklayıcı Değişkenler Varken Ardışık Bağıntının Düzeltilmesi

- Yani ρ'yu biliyorsak, $\tilde{y}_t$'nin $\tilde{x}_t$ üzerine regresyonundan β_0 ve β_1'i tahmin edebiliriz (tahmin edilen sabiti $(1-\rho)$ ile bölmek şartıyla).
- Daha fazla açıklayıcı değişken olması çok fazla şey değiştirmez. $t \geq 2$ için aşağıdaki denklemi kullanabiliriz:

$$\tilde{y}_t = (1-\rho)\beta_0 + \beta_1 \tilde{x}_{t1} + \ldots + \beta_k \tilde{x}_{tk} + e_t, \qquad (26)$$

- Burada $\tilde{x}_{tj} = x_{tj} - \rho x_{t-1,j}$. $t = 1$ için, $\tilde{y_1} = (1-\rho^2)^{0.5} y_1$, $\tilde{x}_{1j} = (1-\rho^2)^{0.5} x_{1j}$ ve sabit $(1-\rho^2)^{0.5}\beta_0$ olacaktır.
- ρ bilindiğinde veriyi dönüştürmek ve OLS uygulamak kolaydır. $\rho = 0$ olmadıkça, GLS tahmincisi, yani dönüştürülmüş veriye OLS tahmincisi, genellikle orjinal OLS tahmincisinden farklı olacaktır. GLS tahmincisi, dönüştürülen denklemin hataları ardışık ilişkisiz ve sabit varyanslı olduğu için, BLUE'dur. Bu denklemin t ve F istatistikleri geçerlidir (en azından asimptotik olarak, ve e_t normal dağılıyorsa daima).

AR(1) Hatalarla Feasible GLS Tahmini

- GLS tahmincisinin eksik yanı, pratikte ρ'nun nadiren bilinmesidir. Ancak, ρ'nun tutarlı bir tahmin edicisini elde edebiliriz. Öncelikle (13)'de olduğu gibi, OLS kalıntılarının gecikmeleri üzerine regresyonunu hesaplamalıyız.
- Sonra, fark-benzeri (quasi-differenced) değişkenleri elde etmek için ρ yerine tahmin edicisi $\hat{\rho}$'yu kullanabiliriz. Sonrasında aşağıdaki denklemi OLS ile tahmin edebiliriz:

$$\tilde{y}_t = \beta_0 \tilde{x}_{t0} + \beta_1 \tilde{x}_{t1} + \ldots + \beta_k \tilde{x}_{tk} + hata_t, \qquad (27)$$

- Burada $t \geq 2$ için $\tilde{x}_{t0} = (1 - \hat{\rho})$ ve $\tilde{x}_{10} = (1 - \hat{\rho})^{0.5}$. Böylece $\hat{\beta}_j$'in **feasible GLS (FGLS)** tahmin edicisini elde etmiş olduk.
- (27)'deki hata terimi e_t'yi içerdiği gibi $\hat{\rho}$'deki tahmin hatasını da içermektedir. $\hat{\rho}$'deki tahmin hatası FGLS tahmin edicilerinin asimptotik dağılımını etkilememektedir.

AR(1) Modelinin Feasible GLS Tahmini

1. y_t'nin $x_{t1}, ..., x_{tk}$ üzerine regresyonundan $t = 1, 2, ..., n$ için OLS kalıntılarını, $\hat{u}_t$, elde edin.
2. (13)'deki regresyonu hesaplayın ve $\hat{\rho}$'i elde edin.
3. $\beta_0, \beta_1, ..., \beta_k$'yı tahmin etmek için (27)'ye OLS uygulayın. Standart hatalar, t ve F istatistikleri asimptotik olarak geçerlidir.

- ρ yerine $\hat{\rho}$'yu kullanmanın getirdiği maliyet ise FGLS tahmin edicisinin istenen sonlu örneklem özelliklerine sahip olmamasıdır. Seriler zayıf bağımlı olduğunda tutarlı olmasına rağmen sapmasız değildir.
- Dahası, (26)'daki e_t normal dağılsa bile, $\hat{\rho}$'deki tahmin hatası yüzünden t ve F istatistikleri ancak asimptotik olarak t ve F dağılımlarına uyar. Bu nedenle küçük örneklemlerde daha dikkatli davranmalıyız.

AR(1) Modelinin Feasible GLS Tahmini

- FGLS tahmin edicisi sapmasız olmadığından, BLUE değildir. Bununla birlikte, ardışık bağıntı için AR(1) modeli geçerliyken, OLS tahmin edicilerinden asimptotik olarak daha etkindir. Halen, zaman serilerinin zayıf bağımlı olduğunu varsayıyoruz.
- ρ'nun tahmin edilmesindeki farklı yöntemlere göre FGLS farklı isimler almaktadır. **Cochrane-Orcutt (CO) tahmini** ilk gözlemi göz ardı eder ve (13)'den $\hat{\rho}$'yu kullanır.
- **Prais-Winsten (PW) tahmini** önceki yolla ilk gözlemi kullanır. Asimptotik olarak ilk gözlemin kullanılıp kullanılmaması sonucu değiştirmez, ancak örneklem boyutu küçükse uygulamada farklılıklar görülebilir.

AR(1) Modelinin Feasible GLS Tahmini

- OLS ve FGLS değişik tahmin yöntemleri olduğundan, aynı tahminleri vermesini bekleyemeyiz. Eğer β_j için benzer tahminler elde ediyorsak, ardışık bağıntı durumunda FGLS tercih edilir. Çünkü daha etkindir ve FGLS test istatistikleri, en azından, asimptotik olarak geçerlidir.
- OLS ve FGLS tahminleri arasında farklar olduğunda, bu farkların istatistiksel olarak anlamlı olup olmadığını saptamak zordur (Hausman (1978) tarafından önerilen yöntem kullanılabilir).
- OLS ve FGLS'in tutarlılık ve asimptotik normalliği, y_t ve x_{jt}'nin zayıf bağımlı olmasına oldukça bağlıdır. Eğer seriler birim kök içeriyorsa bazı sorunlar ortaya çıkabilir ki bunlar Bölüm 18'de ele alınacaktır.

EXAMPLE 12.5
(Static Phillips Curve)

Table 12.2 presents OLS and iterated Cochrane-Orcutt estimates of the static Phillips curve from Example 10.1.

Table 12.2

Dependent Variable: *inf*

Coefficient	OLS	Cochrane-Orcutt
unem	.468 (.289)	−.665 (.320)
intercept	1.424 (1.719)	7.580 (2.379)
$\hat{\rho}$	———	.774 (.091)
Observations *R*-Squared	49 .053	48 .086

- Basit regresyon modelinde

$$y_t = \beta_0 + \beta_1 x_t + u_t, \quad t = \ 1, 2, ..., n \tag{28}$$

- u_t, AR(1) sürecini takip etsin. Bölüm 11.3'de bahsettiğimiz gibi, y_t ve x_t, 1. dereceden bütünleşikse, yani I(1) ise, OLS ile çıkarım yapmak yanıltıcı olacaktır.
- (28)'deki hatalar, $\{u_t\}$, random walk süreci takip ediyorsa, denklem anlamsızdır, çünkü u_t'nin varyansı zamanla birlikte, t ile, büyümektedir. Denklemin farkını almak gerekmektedir:

$$\Delta y_t = \beta_1 \Delta x_t + \Delta u_t, \quad t = \ 2, ..., n \tag{29}$$

- u_t random walk süreci takip ediyorsa, $e_t \equiv \Delta u_t$, sıfır ortalamalı, sabit varyansı ve ardışık olarak ilişkisizdir. Buradan da, e_t ve Δu_t'nin ilişkisiz olduğunu varsayarsak, (29)'u OLS ile tahmin edebiliriz (ilk gözlemi kaybederek).

Otokorelasyona Dirençli Standart Hatalar (Serial Correlation Robust Standard Errors)

- Değişen varyansa dirençli standart hataların nasıl hesaplandığını daha önce görmüştük.
- Benzer şekilde OLS tahmin edicileri betaşapkalar için otokorelasyona dirençli standart hatalar hesaplanabilir.
- Aslında standart hataları hem değişen varyansa hem de otokorelasyona dirençli hale getirmek mümkündür.

Zaman Serilerinde Değişen Varyans

- Değişen varyans problemi genellikle kesit-veri analizlerinde karşımıza çıkar.
- Bu nedenle zaman serilerinde değişen varyans çoğu zaman göz ardı edilen bir problemdir.
- Ancak değişen varyans probleminin zaman serileri modellerinde de karşımıza çıkabileceği unutulmamalıdır.
- Bu nedenle zaman serilerinde değişen varyans probleminin yol açacağı problemleri gözden geçirmek faydalı olabilir.

- TS.1'-TS.2'-TS.3' ve TS.5' varsayımlarının sağlandığını ancak TS.4' varsayımının sağlanmadığını düşünelim.
- TS.1': zayıf bağımlılık ve kovaryans-durağanlık koşulları sağlanıyor.
- TS.2': modelde spesifikasyon hatası, dışlanmış değişkenler ve CEV ölçme hatalarının olmadığı anlamına geliyordu.
- TS.5': hata teriminde otokorelasyon yok.
- TS.4': hata teriminin varyansının sabit olduğunu söyleyen varsayım sağlanmıyor.

- Zaman serileri modellerinde değişen varyans olması durumunda OLS tahmin edicileri, önceki varsayımlar altında, hala tutarlıdır.
- Sapmasızlık varsayımları geçerliyse değişen varyans bu özelliğe de zarar vermez.
- Değişen varyansın varlığı durumunda izlenebilecek bir yok, kesit-veri analizinde olduğu gibi, "değişen varyansa dirençli standart hatalar" kullanmaktır. Diğer test istatistikleri de dirençli hale getirilebilir.
- Çoğu ekonometri programı bu dirençli standart hataları (robust standard errors) otomatik olarak hesaplamaktadır.

- Daha önce Bölüm 9'da gördüğümüz heteroskedasticity testleri zaman serilerinde de kullanılabilir.
- Örneğin Breusch-Pagan testi hesaplanabilir.
- Hatırlarsak bu testin ilk adımında OLS ile model tahmin ediliyor ve kalıntı terimleri hesaplanıyordu.
- İkinci adımda test regresyonu kuruluyordu. Yani kalıntıların karesinin modelde yer alan açıklayıcı değişkenler üzerine regresyon kuruluyordu.
- Daha sonra da bu regresyonun anlamlılığı test ediliyordu (F ya da LM testi ile)
- "Değişen varyans yoktur" boş hipotezi altında eğer anlamlı bir F ya da LM istatistiği bulunursa hata teriminde değişen varyans var demektir.

Zaman Serilerinde Değişen Varyans

- Bu amaçla White testi de kullanılabilir.
- Bu testin adımları aynıdır.
- Modelin heteroskedastik olduğu bulunursa OLS yerine WLS yöntemi kullanılabilir.
- Ancak yan regresyonun hata teriminin otokorelasyonlu olması durumunda bu testler geçersiz olur.
- Zaman serilerinde bu durumla oldukça sık karşılaşılır.
- Hata varyansının zaman içindeki değişimi dinamik bir yapıya sahipse buna uygun modellerin kullanılması gerekir.
- Özellikle finansal zaman serilerinde oynaklığın zaman içindeki davranışı otoregresif bir yapıya sahiptir.

$$return_t = \beta_0 + \beta_1 return_{t-1} + u_t$$

- $t_{\beta_1} = 1.55$. EMH aleyhinde kanıt olmadığı görünüyor???
- BP testi için yardımcı regresyon

$$\widehat{\mathrm{u}}_t^2 = \underset{(0.043)}{4.66} - \underset{(0.201)}{1.104}\, return_{t-1} + residual_t$$

$$n = 689 \quad R^2 = 0.042$$

- $t_{return_{t-1}} = -5.5$. Değişen varyans yönünde güçlü kanıt vardır.

Zaman Serilerinde Değişen Varyans Örnek: EMH

$$\widehat{u}_t^2 = \underset{(0.043)}{4.66} - \underset{(0.201)}{1.104}\, return_{t-1} + residual_t$$

$$n = 689 \quad R^2 = 0.042$$

- Bu sonuçlara göre bir önceki dönemdeki hisse senedi getirisi yüksekse cari dönemdeki oynaklık (volatility) daha düşük olmaktadır.
- Getirinin beklenen değeri bir önceki döneme bağlı olmasa da (EMH t-testi buna işaret ediyor) varyansın getiriyle birlikte hareket ettiği görülmektedir.
- Bu birçok finansal getiri serisinde karşılaşılan bir özelliktir.

Otoregressif Koşullu Değişen Varyans(Autoregressive Conditional Heteroskedasticity-ARCH)

- Basit regresyon modelini düşünelim:

$$y_t = \beta_0 + \beta_1 z_t + u_t$$

- Hata terimi için aşağıdaki modelin geçerli olduğunu düşünelim:

$$E(u_t^2 | u_{t-1}, u_{t-2}, ...) = E(u_t^2 | u_{t-1}) = \alpha_0 + \alpha_1 u_{t-1}^2$$

$$u_t^2 = \alpha_0 + \alpha_1 u_{t-1}^2 + v_t \quad ARCH(1) \quad Modeli$$

- $\alpha_0 > 0$, $\alpha_1 \geq 0$, $\alpha_1 < 0$

Otoregressif Koşullu Değişen Varyans(Autoregressive Conditional Heteroskedasticity-ARCH)

- Eğer Gauss-Markov varsayımları sağlanıyorsa ARCH hata terimleri altında OLS sapmasız ve tutarlı bir tahmin çerçevesi sunar.
- Statik ve dağıtılmış gecikme modellerinde TS' varsayımları sağlanıyorsa OLS tutarlıdır. Peki bu durumda ARCH etkisi ile ilgilenmemize gerek var mı?
- Evet. İki nedenden dolayı. Birincisi, OLS'den daha etkin tahmin ediciler bulmak mümkündür, e.g. WLS (Asimptotik olarak)
- İkincisi, ARCH modeli varyansın (oynaklığın) zaman içindeki davranışını betimlemesi açısından önemlidir. Örneğin, finansal zaman serilerinde oynaklığın yüksek olduğu dönemleri yine yüksek oynaklık dönemlerinin takip ettiği görülmektedir.

Otoregressif Koşullu Değişen Varyans(Autoregressive Conditional Heteroskedasticity-ARCH) Örnek

- Hisse senedi getirileri için ARCH(1) modeli tahmini

$$\widehat{\mathrm{u}}_t^2 = \underset{(0.44)}{+2.95} + \underset{(0.036)}{0.337}\,\hat{u}_{t-1}^2 + residual_t$$
$$n = 688 \quad R^2 = 0.114$$

- Yüksek t istatistiği güçlü ARCH etkisinin varlığına işaret etmektedir.