



Ανάλυση Δεδομένων στη Λογιστική και Χρηματοοικονομική

Ενότητα 7^η

Γραμμικό Υποδείγμα Παλινδρόμησης – Εμπειρικές εφαρμογές

Περιγραφή Ενότητας

- Εμπειρικά παραδείγματα
- Στρατηγικές για την ανάπτυξη οικονομετρικών μοντέλων
- Διαδικασία απόφασης για ανάλυση πολλαπλού γραμμικό υποδείγματος παλινδρόμησης

Εμπειρικό παράδειγμα Ι

Απλό γραμμικό υπόδειγμα παλινδρόμησης

Ας ξεκινήσουμε με το απλό γραμμικό υπόδειγμα παλινδρόμησης

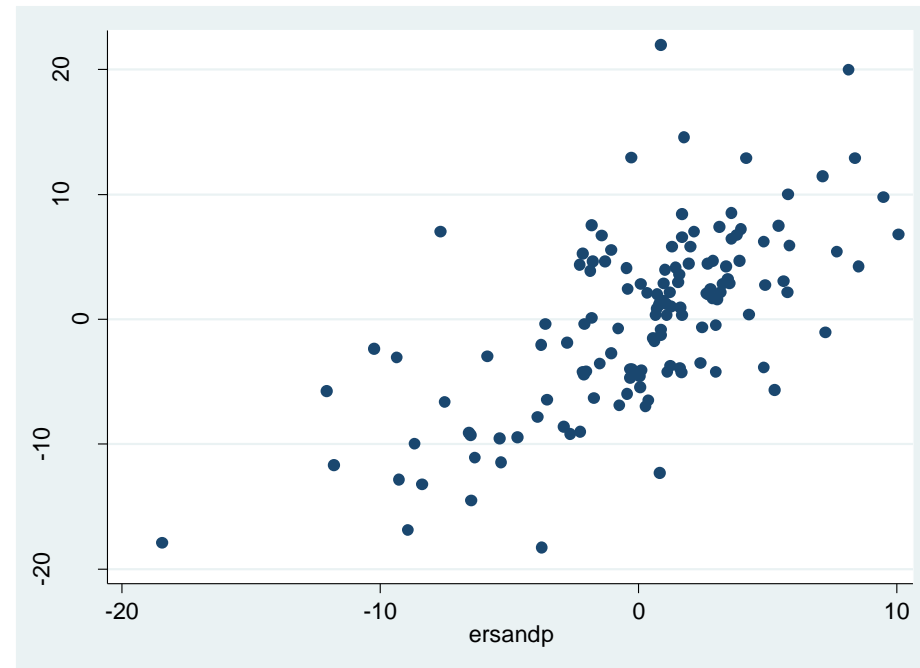
$$(R_{\text{MICROSOFT}} - r_f)_t = \beta_0 + \beta_1(R_M - r_f)_t + u_t$$

όπου R_F η απόδοση της μετοχής της *Microsoft*, r_f το επιτόκιο μηδενικού κινδύνου και R_M η απόδοση της μετοχής του αμερικανικού χρηματιστηριακού δείκτη S&P500.

- Περίοδος: 29/01/1988- 31/10/2012

Εκτιμάμε το μοντέλο παλινδρόμησης με τη χρήση του STATA και έχουμε :

- **Scatter Plot**



- Ερμηνεία

```
. regress ermsoft ersand
```

Source	SS	df	MS	Number of obs =	325
Model	11968.8168	1	11968.8168	F(1, 323) =	74.18
Residual	52117.0991	323	161.353248	Prob > F =	0.0000
Total	64085.9159	324	197.796037	R-squared =	0.1868
				Adj R-squared =	0.1842
				Root MSE =	12.702

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ersandp	1.325376	.1538871	8.61	0.000	1.022628 1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211 .7748094

Any set of (X, Y) are specific to the sample used in their estimation. It would be desirable to have an idea of how reliable/precise these estimates are. Thus we need an estimate given by the **standard error**.

The standard error

$$SE(\hat{\alpha})$$

The standard error

$$SE(\hat{\beta})$$

The beta coefficient (the slope coefficient) estimate is 1.33

$$\hat{\beta} = \frac{Cov(ExcR_{Microsoft}, ExcR_{SP500})}{Var(ExcR_{SP500})}$$

The alpha coefficient estimate is -0.6137

$$\hat{\alpha} = E(Y) - \hat{\beta} E(X)$$

- Έλεγχος υποθέσεων – Προσέγγιση κριτικής τιμής (Critical value approach)

Two –sided Test

$$H_0 : \alpha = 0$$

$$H_A : \alpha \neq 0$$

Critical value approach

$$test\ statistic = \frac{\hat{a} - a}{SE(\hat{a})}$$

We do **not** reject the Null Hypothesis for a ; thus a is **insignificant**

$$H_0 : \beta = 0$$

$$H_A : \beta \neq 0$$

a = 5% significance level

$$test\ statistic = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})}$$

We reject the Null Hypothesis for b ; thus b is **significant**

. regress ermsoft ersand

Source	SS	df	MS	Number of obs =	325
Model	11968.8168	1	11968.8168	F(1, 323) =	74.18
Residual	52117.0991	323	161.353248	Prob > F =	0.0000
Total	64085.9159	324	197.796037	R-squared =	0.1868
				Adj R-squared =	0.1842
				Root MSE =	12.702

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ersandp	1.325376	.1538871	8.61	0.000	1.022628 1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211 .7748094

- Έλεγχος υποθέσεων – Προσέγγιση με p-value (p-value approach)

Two –sided Test

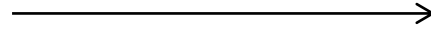
$$H_0 : \alpha = 0$$

$$H_A : \alpha \neq 0$$

$$H_0 : \beta = 0$$

$$H_A : \beta \neq 0$$

P-value approach



p-value is termed as the “plausibility” of the Null Hypothesis; the smaller the p-value, the less plausible is the null hypothesis. Is the largest significance level at which we fail to reject the null hypothesis.

```
. regress ermsoft ersand
```

Source	SS	df	MS			
Model	11968.8168	1	11968.8168	Number of obs =	325	
Residual	52117.0991	323	161.353248	F(1, 323) =	74.18	
Total	64085.9159	324	197.796037	Prob > F =	0.0000	
				R-squared =	0.1868	
				Adj R-squared =	0.1842	
				Root MSE =	12.702	

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ersandp	1.325376	.1538871	8.61	0.000	1.022628	1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211	.7748094

- Έλεγχος υποθέσεων – Προσέγγιση με διαστήματα εμπιστοσύνης (Confidence interval approach)

Two –sided Test

$$H_0 : \alpha = 0$$

$$H_A : \alpha \neq 0$$

$$H_0 : \beta = 0$$

$$H_A : \beta \neq 0$$

Confidence interval approach

a = 5% significance level

$$\hat{a} \pm t_{crit} SE(\hat{a})$$

$$\hat{\beta} \pm t_{crit} SE(\hat{\beta})$$

(-0.002,0.014)

We do **not** reject the Null Hypothesis for a ; thus a is **Insignificant**, since 0 lies within confidence interval

(1.05,1.45)

We reject the Null Hypothesis for b ; thus b is **significant**, since 0 does **not** lie within confidence interval

```
. regress ermsoft ersand
```

Source	SS	df	MS	Number of obs =	325
Model	11968.8168	1	11968.8168	F(1, 323) =	74.18
Residual	52117.0991	323	161.353248	Prob > F =	0.0000
Total	64085.9159	324	197.796037	R-squared =	0.1868
				Adj R-squared =	0.1842
				Root MSE =	12.702

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ersandp	1.325376	.1538871	8.61	0.000	1.022628 1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211 .7748094

Έλεγχος υποθέσης με Wald test

Μηδενική υπόθεση $\beta_1 = 1$

Εναλλακτική υπόθεση $\beta_1 \neq 1$

```
. test (ersandp=1)

( 1)  ersandp = 1

           F( 1, 133) =    0.00
           Prob > F =    0.9616
```

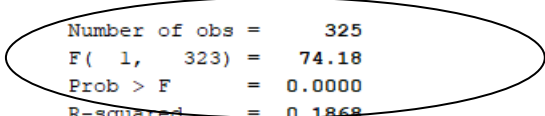
- ❑ $F(1,133)$: F-statistic with one restriction and $T-k=135-2=133$
- ❑ We do not reject the Null hypothesis since the $p\text{-value} > 0.05$ (significant level = 5%)

Έλεγχος υποθέσεων

```
. regress ermsoft ersand
```

Source	SS	df	MS			
Model	11968.8168	1	11968.8168	Number of obs =	325	
Residual	52117.0991	323	161.353248	F(1, 323) =	74.18	
Total	64085.9159	324	197.796037	Prob > F =	0.0000	
				R-squared =	0.1868	
				Adj R-squared =	0.1842	
				Root MSE =	12.702	

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ersandp	1.325376	.1538871	8.61	0.000	1.022628	1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211	.7748094



$$H_0 : \beta = 0$$

$$H_A : \beta \neq 0$$

We reject the Null Hypothesis for levels of significance 1%, 5% and 10%, since p-value of F – statistic is 0.000. In this case F-test statistic is equal to the square of the slope t-stat.

```
. test ( ersandp=1) (_cons=1)
```

```
( 1) ersandp = 1
```

```
( 2) _cons = 1
```

```
F( 2, 323) = 4.59
Prob > F = 0.0109
```

$$H_0 : \alpha = 1 \quad \text{and} \quad \beta = 1$$

$$H_A : \alpha \neq 1 \quad \text{or} \quad \beta \neq 1$$

We reject the joint Null Hypothesis for levels of significance 5% and 10%.

□ The F-version is adjusted for small sample bias and should be used when the regression is estimated using a small sample

• Πολλαπλό γραμμικό υπόδειγμα παλινδρόμησης

Συνεχίζουμε με το πολλαπλό γραμμικό υπόδειγμα παλινδρόμησης και ορίζεται ως εξής:

$$(R_{Microsoft} - r_f)_t = \alpha + \beta_1(R_M - r_f)_t + \beta_2(d_{prod})_t + \beta_3(d_{credit})_t + \beta_4(d_{inflation})_t + \beta_5(d_{money})_t + \beta_6(d_{spread})_t + \beta_7(r_{term})_t + u_t$$

Εκτιμάμε το μοντέλο παλινδρόμησης με τη χρήση του STATA και έχουμε :

```
. regress ermsoft ersand dprod dcredit dinflation dmoney dspread rterm
```

Source	SS	df	MS	Number of obs = 324		
Model	13202.4359	7	1886.06227	F(7, 316) =	11.77	
Residual	50637.6544	316	160.245742	Prob > F =	0.0000	
Total	63840.0903	323	197.647338	R-squared =	0.2068	
				Adj R-squared =	0.1892	
				Root MSE =	12.659	

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ersandp	1.360448	.1566147	8.69	0.000	1.052308	1.668587
dprod	-1.425779	1.324467	-1.08	0.283	-4.031668	1.180109
dcredit	-.0000405	.0000764	-0.53	0.596	-.0001909	.0001098
dinflation	2.95991	2.166209	1.37	0.173	-1.302104	7.221925
dmoney	-.0110867	.0351754	-0.32	0.753	-.0802944	.0581209
dspread	5.366629	6.913915	0.78	0.438	-8.236496	18.96975
rterm	4.315813	2.515179	1.72	0.087	-.6327998	9.264426
_cons	-.1514086	.9047867	-0.17	0.867	-1.931576	1.628759

Έλεγχος πολλαπλών υποθέσεων

$H_0: \beta_2 = 0 \text{ and } \beta_3 = 0 \text{ and } \beta_4 = 0 \text{ and } \beta_5 = 0 \text{ and } \beta_6 = 0$
 $H_A: \beta_2 \neq 0 \text{ or } \beta_3 \neq 0 \text{ or } \beta_4 \neq 0 \text{ or } \beta_5 \neq 0 \text{ or } \beta_6 \neq 0$

```
. test (dprod dcredit dinflation dmoney dsread)

( 1)  dprod = 0
( 2)  dcredit = 0
( 3)  dinflation = 0
( 4)  dmoney = 0
( 5)  dsread = 0

          F( 5, 316) = 0.85
          Prob > F = 0.5131
```

Σύγκριση Απλού και Πολλαπλού γραμμικού υποδείγματος

$$(R_{Microsoft} - r_f)_t = \gamma_0 + \gamma_1(R_M - r_f)_t + \varepsilon_t$$

$$(R_{Microsoft} - r_f)_t = \alpha + \beta_1(R_M - r_f)_t + \beta_2(d_{prod})_t + \beta_3(d_{credit})_t + \beta_4(d_{inflation})_t + \beta_5(d_{money})_t + \beta_6(d_{spread})_t + \beta_7(r_{term})_t + u_t$$

. regress ermsoft ersand

Source	SS	df	MS	Number of obs = 325
Model	11968.8168	1	11968.8168	F(1, 323) = 74.18
Residual	52117.0991	323	161.353248	Prob > F = 0.0000
Total	64085.9159	324	197.796037	R-squared = 0.1868
				Adj R-squared = 0.1842
				Root MSE = 12.702

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ersandp	1.325376	.1538871	8.61	0.000	1.022628 1.628124
_cons	-.6137005	.705782	-0.87	0.385	-2.002211 .7748094

. regress ermsoft ersand dprod dcredit dinflation dmoney dspread rterm

Source	SS	df	MS	Number of obs = 324
Model	13202.4359	7	1886.06227	F(7, 316) = 11.77
Residual	50637.6544	316	160.245742	Prob > F = 0.0000
Total	63840.0903	323	197.647338	R-squared = 0.2068
				Adj R-squared = 0.1892
				Root MSE = 12.659

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]
ersandp	1.360448	.1566147	8.69	0.000	1.052308 1.668587
dprod	-1.425779	1.324467	-1.08	0.283	-4.031668 1.180109
dcredit	-.0000405	.0000764	-0.53	0.596	-.0001909 .0001098
dinflation	2.95991	2.166209	1.37	0.173	-1.302104 7.221925
dmoney	-.0110867	.0351754	-0.32	0.753	-.0802944 .0581209
dspread	5.366629	6.913915	0.78	0.438	-8.236496 18.96975
rterm	4.315813	2.515179	1.72	0.087	-.6327998 9.264426
_cons	-.1514086	.9047867	-0.17	0.867	-1.931576 1.628759

- Συνεχίζουμε την ανάλυσή μας με τις υποθέσεις, τους διαγνωστικούς ελέγχους και την αντιμετώπιση σε περίπτωση παραβίασης των υποθέσεων.

**Υπόθεση 1: Ο διαταρακτικός όρος
έχει μέσο όρο μηδέν**

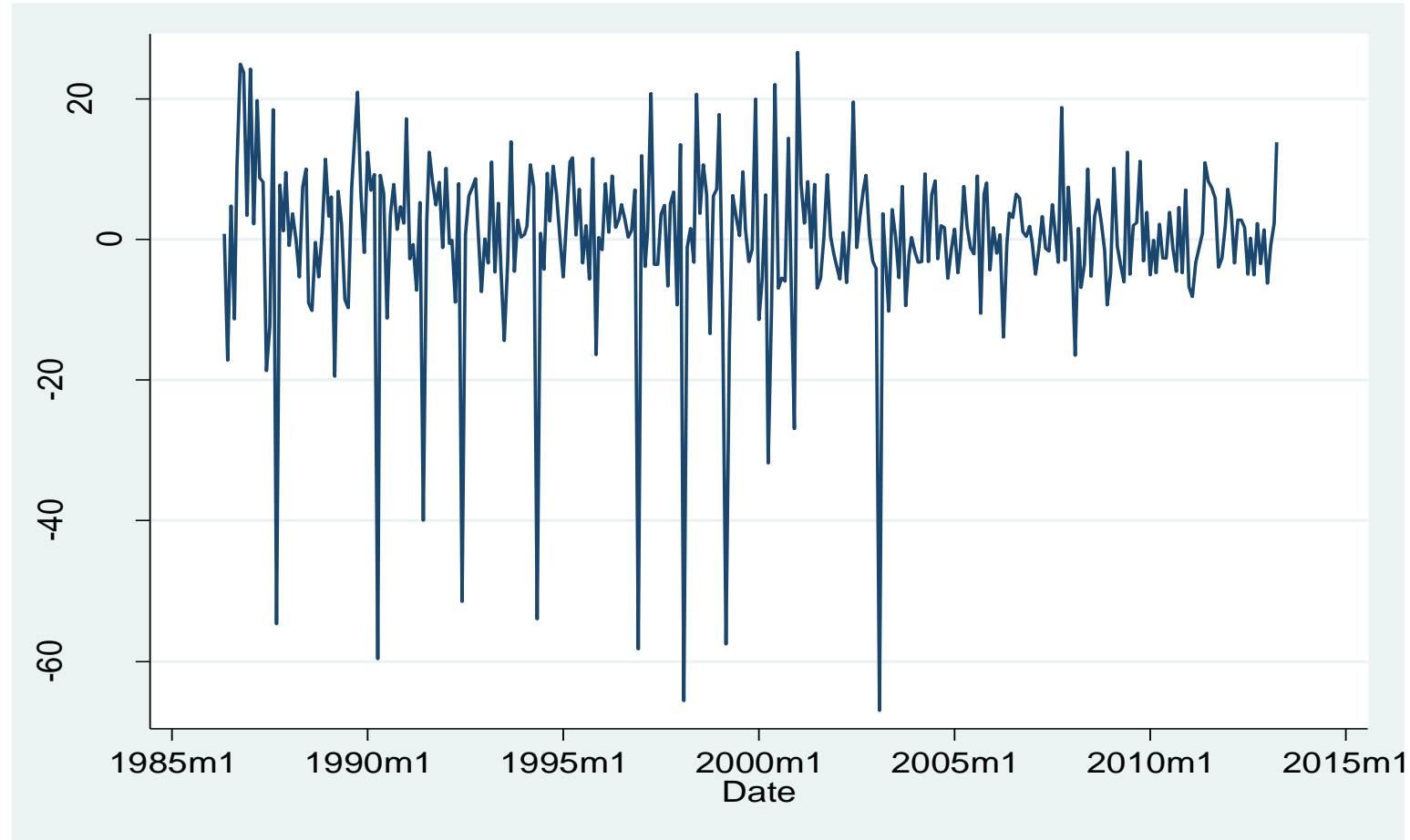
- Ο μέσος όρος των καταλοίπων θα είναι πάντα μηδενικός υπό τον όρο ότι υπάρχει σταθερός όρος στην παλινδρόμηση.

Υπόθεση 2: Ομοσκεδαστικότητα

$$\text{Var}(u_t) = \sigma^2 < \infty$$

Διάγνωση της Ετεροσκεδαστικότητας

- Γραφικός έλεγχος



H_0 : Ομοσκεδαστικότητα (*Homoskedasticity*)
 H_A : Ετεροσκεδαστικότητα (*Heteroskedasticity*)

```
. estat imtest, white
```

```
White's test for Ho: homoskedasticity  
against Ha: unrestricted heteroskedasticity
```

```
chi2(35)      =    11.12  
Prob > chi2   =    1.0000
```

```
Cameron & Trivedi's decomposition of IM-test
```

Source	chi2	df	p
Heteroskedasticity	11.12	35	1.0000
Skewness	10.26	7	0.1742
Kurtosis	8.86	1	0.0029
Total	30.24	43	0.9289

White standard errors

```
. regress ermsoft ersand dprod dcredit dinflation dmoney dspread rterm, vce(robust)
```

Linear regression

Number of obs = 324
F(7, 316) = 14.87
Prob > F = 0.0000
R-squared = 0.2068
Root MSE = 12.659

ermsoft	Robust		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
ersandp	1.360448	.145839	9.33	0.000	1.07351	1.647386
dprod	-1.425779	.8630263	-1.65	0.100	-3.123783	.2722243
dcredit	-.0000405	.0000544	-0.75	0.456	-.0001475	.0000664
dinflation	2.95991	1.786173	1.66	0.098	-.554385	6.474206
dmoney	-.0110867	.0274214	-0.40	0.686	-.0650384	.0428649
dspread	5.366629	4.630536	1.16	0.247	-3.74395	14.47721
rterm	4.315813	2.149673	2.01	0.046	.0863325	8.545294
_cons	-.1514086	.8089487	-0.19	0.852	-1.743015	1.440198

Υπόθεση 3: Μη αυτοσυσχέτιση
μεταξύ των διαταρακτικών όρων

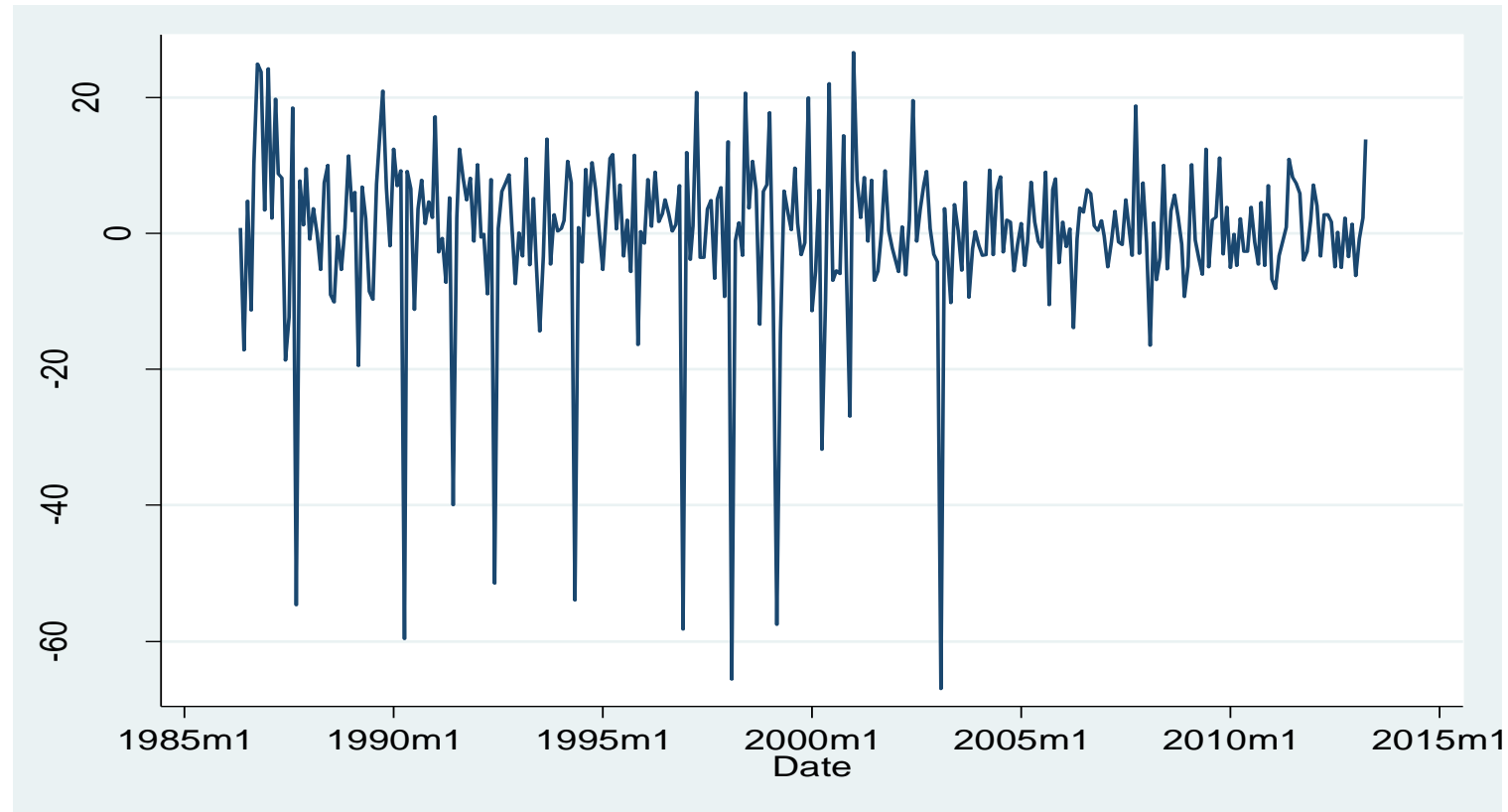
$$\text{Cov}(u_i, u_j) = 0$$

Διάγνωση της Αυτοσυσχέτισης

- Γραφικός έλεγχος

A. με βάση το διάγραμμα διασποράς των καταλοίπων. Συγκεκριμένα παρατηρούμε τη γραφική παράσταση των καταλοίπων u_{t-1} και u_t .

B. με βάση ένα διάγραμμα χρονοσειράς, όπου στον οριζόντιο άξονα είναι ο χρόνος και στον κατακόρυφο άξονα οι τιμές των καταλοίπων



- Ο έλεγχος Durbin-Watson (DW)

$$DW \approx 2(1 - \hat{\rho})$$

Η στατιστική Durbin-Watson είναι 2.17, κοντά στο 2

```
. estat dwatson
```

```
Durbin-Watson d-statistic( 8, 324) = 2.165384
```

$$u_t = \rho u_{t-1} + v_t$$

$H_0: \rho = 0$ (Μη αυτοσυσχέτιση)

$H_A: \rho \neq 0$ (Αυτοσυσχέτιση)

- Ο έλεγχος Breusch-Godfrey(BG)

$H_0: \rho_1=0$ και $\rho_2=0$ και ... $\rho_{12}=0$
 $H_A: \rho_1 \neq 0$ ή $\rho_2 \neq 0$ ή... $\rho_{12} \neq 0$

```
. estat bgodfrey, lags (12)
```

Breusch-Godfrey LM test for autocorrelation

lags(p)	chi2	df	Prob > chi2
12	25.974	12	0.0108

H0: no serial correlation

p-value < 0.05,
άρα απορρίπτουμε τη μηδενική
υπόθεση της μη αυτοσυσχέτισης των
καταλοίπων

Καθορίστε τον αριθμό των lags ίσο με 12.
Δεν υπάρχει προφανής απάντηση σε αυτό,
μπορείτε να πειραματιστείτε σε μια σειρά από
αριθμούς. Μπορείτε να συσχετίσετε τον αριθμό
των lags με τη συχνότητα των δεδομένων σας.
(για μηνιαία δεδομένα χρησιμοποιήστε 12, για
τριμηνιαία δεδομένα 4, κ.λπ.)

Μια Πιο «Πρακτική» Προσέγγιση στην Αντιμετώπιση της Ετεροσκεδαστικότητας Αυτοσυσχέτισης

- Οι **Newey West** αναπτύσσουν έναν εκτιμητή διακυμάνσεων-συνδυακυμάνσεων των συντελεστών (δηλαδή υπολογίζουν τις τυπικές εκτιμήσεις των σφαλμάτων των συντελεστών) που να είναι συνεπής **ΚΑΙ** στην ετεροσκεδαστικότητα **ΚΑΙ** αυτοσυσχέτιση.
- Εκτιμάμε το μοντέλο παλινδρόμησης με τη χρήση του STATA χρησιμοποιώντας τη Newey West μεθοδολογία και έχουμε :

$$m(T) = \text{floor}[4(T/100)^{2/9}] .$$

```
. newey ermssoft ersand dprod dcredit dinflation dmoney dspread rterm, lag(5)
```

```
Regression with Newey-West standard errors      Number of obs =      324
maximum lag: 5                                F( 7, 316) =      14.85
                                              Prob > F      =      0.0000
```

ermssoft	Coef.	Newey-West Std. Err.	t	P> t	[95% Conf. Interval]	
ersandp	1.360448	.1468806	9.26	0.000	1.07146	1.649435
dprod	-1.425779	.7693381	-1.85	0.065	-2.939452	.0878929
dcredit	-.0000405	.0000496	-0.82	0.414	-.0001381	.000057
dinflation	2.95991	1.971965	1.50	0.134	-.9199292	6.83975
dmoney	-.0110867	.0292309	-0.38	0.705	-.0685985	.0464251
dspread	5.366629	4.46252	1.20	0.230	-3.413378	14.14664
rterm	4.315813	2.248346	1.92	0.056	-.1078064	8.739433
_cons	-.1514086	.7402347	-0.20	0.838	-1.60782	1.305003

**Newey & West για
ετεροσκεδαστικότητα
ΚΑΙ αυτοσυσχέτιση.**

Σύγκριση των αποτελεσμάτων πριν και μετά τη χρήση των τυπικών σφαλμάτων συνεπή σε ετεροσκεδαστικότητα και αυτοσυσχέτιση.

Μετά....

```
. newey ermsoft ersand dprod dcredit dinflation dmoney dspread rterm, lag(5)

Regression with Newey-West standard errors      Number of obs =      324
maximum lag: 5                                F( 7, 316) =      14.85
                                              Prob > F       =      0.0000
```

ermsoft	Newey-West		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
ersandp	1.360448	.1468806	9.26	0.000	1.07146	1.649435
dprod	-1.425779	.7693381	-1.85	0.065	-2.939452	.0878929
dcredit	-.0000405	.0000496	-0.82	0.414	-.0001381	.000057
dinflation	2.95991	1.971965	1.50	0.134	-.9199292	6.83975
dmoney	-.0110867	.0292309	-0.38	0.705	-.0685985	.0464251
dspread	5.366629	4.46252	1.20	0.230	-3.413378	14.14664
rterm	4.315813	2.248346	1.92	0.056	-.1078064	8.739433
_cons	-.1514086	.7402347	-0.20	0.838	-1.60782	1.305003

Πριν....

```
. regress ermsoft ersand dprod dcredit dinflation dmoney dspread rterm
```

Source	SS	df	MS	Number of obs =	324
Model	13202.4359	7	1886.06227	F(7, 316) =	11.77
Residual	50637.6544	316	160.245742	Prob > F	= 0.0000
Total	63840.0903	323	197.647338	R-squared	= 0.2068
				Adj R-squared	= 0.1892
				Root MSE	= 12.659

ermsoft	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ersandp	1.360448	.1566147	8.69	0.000	1.052308	1.668587
dprod	-1.425779	1.324467	-1.08	0.283	-4.031668	1.180109
dcredit	-.0000405	.0000764	-0.53	0.596	-.0001909	.0001098
dinflation	2.95991	2.166209	1.37	0.173	-1.302104	7.221925
dmoney	-.0110867	.0351754	-0.32	0.753	-.0802944	.0581209
dspread	5.366629	6.913915	0.78	0.438	-8.236496	18.96975
rterm	4.315813	2.515179	1.72	0.087	-.6327998	9.264426
_cons	-.1514086	.9047867	-0.17	0.867	-1.931576	1.628759

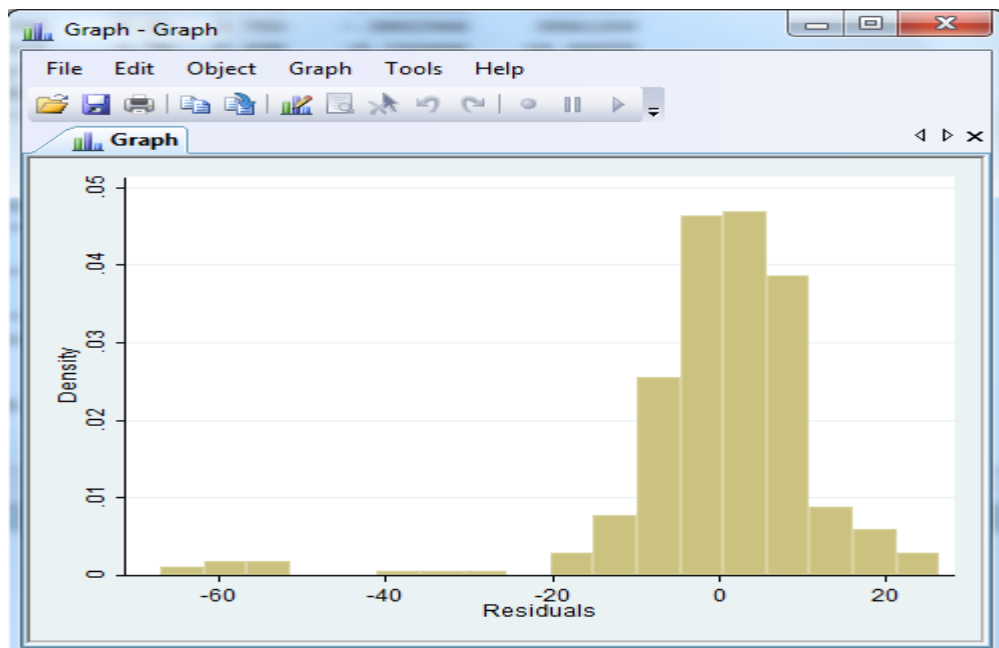
Υπόθεση 4: Οι ανεξάρτητες μεταβλητές είναι μη στοχαστικές

$$\text{Cov}(u_t, x_t) = 0$$

Υπόθεση 5: Ο διαταρακτικός όρος
κατανέμεται κανονικά με μέσο
μηδέν και σταθερή διακύμανση.

$$u_t \sim N(0, \sigma^2)$$

Αν η τιμή της στατιστικής των Jarque – Bera (JB) είναι μεγαλύτερη από την κριτική τιμή της χ^2 κατανομής (τιμή των πινάκων) τότε απορρίπτουμε τη μηδενική υπόθεση οπότε τα κατάλοιπα δεν κατανέμονται κανονικά. Αντίθετα, αν η τιμή της στατιστικής των Jarque – Bera (JB) είναι μικρότερη από την κριτική τιμή της χ^2 κατανομής (τιμή των πινάκων) τότε η μηδενική υπόθεση δεν απορρίπτεται και συνεπώς τα κατάλοιπα κατανέμονται κανονικά.



H_0 : Κανονικότητα (Normality)
 H_A : Μη Κανονικότητα (Non-Normality)

p-value is 0.000, άρα απορρίπτουμε την μηδενική υπόθεση ότι τα κατάλοιπα ακολουθούν την κανονική κατανομή

Skewness/Kurtosis tests for Normality

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2 (2)	Prob>chi2
resid	324	0.0000	0.0000	.	0.0000

Μια έμμεση (implicit) υπόθεση: Οι ανεξάρτητες μεταβλητές δε συσχετίζονται μεταξύ τους

1. Ο ευκολότερος τρόπος για να μετρήσετε την έκταση της πολυγραμμικότητας είναι απλώς να εξετάσετε τη μήτρα συσχετίσεων μεταξύ των ανεξάρτητων μεταβλητών

Πίνακας Συσχετίσεων (Correlation Matrix)

```
. correlate ersand dprod dcredit dinflation dmoney dspread rterm
(obs=324)
```

	ersandp	dprod	dcredit	dinfla~n	dmoney	dspread	rterm
ersandp	1.0000						
dprod	-0.0253	1.0000					
dcredit	0.0364	0.1411	1.0000				
dinflation	-0.0038	-0.1243	0.0452	1.0000			
dmoney	0.0241	-0.1301	-0.0117	-0.0980	1.0000		
dspread	-0.1758	-0.0556	0.0153	-0.2248	0.2136	1.0000	
rterm	-0.0220	-0.0024	0.0097	-0.0542	-0.0862	0.0016	1.0000

Κανένα πρόβλημα πολυγραμμικότητας, αφού ο συσχετισμός μεταξύ των ερμηνευτικών μεταβλητών είναι σχετικά μικρός

2. Ένα δεύτερος τρόπος είναι να προσθέσετε ή να αφαιρέσετε ερμηνευτικές μεταβλητές, και εξετάστε αν υπάρξει μια μεγάλη αλλαγή στις τιμές των συντελεστών

Π.χ. Αφαιρέσετε το d dspread από το υπόδειγμα παλινδρόμησης

Χωρίς d dspread

```
. newey ermsoft ersandp dprod dcredit dinflation dmoney dspread rterm,lag(5)
Regression with Newey-West standard errors      Number of obs =      324
maximum lag: 5                                F( 7, 316) =      14.85
                                              Prob > F      =      0.0000
```

ermsoft	Newey-West		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
ersandp	1.360448	.1468806	9.26	0.000	1.07146	1.649435
dprod	-1.425779	.7693381	-1.85	0.065	-2.939452	.0878929
dcredit	-.0000405	.0000496	-0.82	0.414	-.0001381	.000057
dinflation	2.95991	1.971965	1.50	0.134	-.9199292	6.83975
dmoney	-.0110867	.0292309	-0.38	0.705	-.0685985	.0464251
dspread	5.366629	4.46252	1.20	0.230	-3.413378	14.14664
rterm	4.315813	2.248346	1.92	0.056	-.1078064	8.739433
_cons	-.1514086	.7402347	-0.20	0.838	-1.60782	1.305003

Με dspread

```
. newey ermsoft ersandp dprod dcredit dinflation dmoney dspread rterm, lag(5)
Regression with Newey-West standard errors      Number of obs =      324
maximum lag: 5                                F( 7, 316) =      14.85
                                              Prob > F      =      0.0000
```

ermsoft	Newey-West		t	P> t	[95% Conf. Interval]	
	Coef.	Std. Err.				
ersandp	1.360448	.1468806	9.26	0.000	1.07146	1.649435
dprod	-1.425779	.7693381	-1.85	0.065	-2.939452	.0878929
dcredit	-.0000405	.0000496	-0.82	0.414	-.0001381	.000057
dinflation	2.95991	1.971965	1.50	0.134	-.9199292	6.83975
dmoney	-.0110867	.0292309	-0.38	0.705	-.0685985	.0464251
dspread	5.366629	4.46252	1.20	0.230	-3.413378	14.14664
rterm	4.315813	2.248346	1.92	0.056	-.1078064	8.739433
_cons	-.1514086	.7402347	-0.20	0.838	-1.60782	1.305003

Χωρίς μεγάλες αλλαγές στις τιμές των συντελεστών, συνεπώς δεν υπάρχει παρουσία πολυσυγγραμικότητας

Διαδικασία απόφασης για ανάλυση πολλαπλού γραμμικό υποδείγματος παλινδρόμησης

Στάδιο 1: Στόχοι της πολλαπλής παλινδρόμησης

- Προβλήματα έρευνας κατάλληλα για πολλαπλή παλινδρόμηση
- Πρόβλεψη με πολλαπλή παλινδρόμηση
- Επεξήγηση με πολλαπλή παλινδρόμηση
- Καθορισμός στατιστικής σχέσης
- Επιλογή εξαρτημένων και ανεξάρτητων μεταβλητών

Στάδιο 2: Έρευνα σχεδιασμού μιας ανάλυσης πολλαπλής παλινδρόμησης

- Το μέγεθος του δείγματος
- Στατιστική ισχύς και μέγεθος δείγματος
- Δημιουργία πρόσθετων μεταβλητών
- Ενσωμάτωση μη μετρικών δεδομένων με ψευδομεταβλητές (Dummy Variables)

Στάδιο 3: Υποθέσεις στην ανάλυση πολλαπλής παλινδρόμησης

- Αξιολόγηση μεμονωμένων μεταβλητών
- Γραμμικότητα του Φαινομένου
- Σταθερή διακύμανση του διαταρακτικού όρου
- Ανεξαρτησία των διαταρακτικών όρων
- Κανονικότητα της κατανομής του διαταρακτικού όρου

Στάδιο 4: Εκτίμηση του μοντέλου παλινδρόμησης και εκτίμηση της συνολικής προσαρμογής

- Εξετάζοντας τη στατιστική σημαντικότητα του μοντέλου μας
- Σημαντικότητα του Γενικού Μοντέλου: Ο Συντελεστής Προσδιορισμού
- Έλεγχοι στατιστικής σημαντικότητας των συντελεστών παλινδρόμησης

Στάδιο 5: Ερμηνεία της παραλλαγής παλινδρόμησης

- Χρησιμοποιώντας τους Συντελεστές Παλινδρόμησης

Στάδιο 6: Επικύρωση των αποτελεσμάτων

Στρατηγικές για την ανάπτυξη οικονομετρικών μοντέλων

Στόχος μας:

Να οικοδομήσουμε ένα στατιστικά επαρκές εμπειρικό μοντέλο το οποίο

- να ικανοποιεί τις παραδοχές του γραμμικού υποδείγματος παλινδρόμησης.
- να είναι περιεκτικό (parsimonious).
- να έχει την κατάλληλη θεωρητική ερμηνεία.
- να έχει το σωστό «σχήμα» - δηλ.
 - ❖ όλα τα πρόσημα στους συντελεστές να είναι «σωστά»
 - ❖ όλα τα μεγέθη των συντελεστών να είναι «σωστά»
- να είναι σε θέση να εξηγήσει τα αποτελέσματα όλων των ανταγωνιστικών μοντέλων

Προσεγγίσεις στην ανάπτυξη οικονομετρικών μοντέλων

Υπάρχουν 2 δημοφιλείς φιλοσοφίες για την ανάπτυξη οικονομετρικών μοντέλων: οι προσεγγίσεις «από ειδικό στο γενικό» (“specific-to-general”) και «από γενικό στο ειδικό» (“general-to-specific”) :

- Το "Specific-to-general" χρησιμοποιήθηκε σχεδόν καθολικά μέχρι τα μέσα της δεκαετίας του 1980 και αφορούσε την έναρξη με το απλούστερο μοντέλο και τη σταδιακή προσθήκη σε αυτό.
- Πραγματοποιήθηκε ελάχιστος διαγνωστικός έλεγχος. Αυτό όμως σήμαινε ότι όλα τα συμπεράσματα ήταν δυνητικά άκυρα.
- Μια εναλλακτική και πιο σύγχρονη προσέγγιση στο πρότυπο κτίριο είναι η μεθοδολογία "LSE" ή Hendry “general-to-specific” .
- Τα πλεονεκτήματα αυτής της προσέγγισης είναι ότι είναι στατιστικά λογική και επίσης η θεωρία στην οποία βασίζονται τα μοντέλα συνήθως δεν έχει τίποτα να πει για τη δομή υστέρησης ενός μοντέλου.

Η προσέγγιση General-to-Specific

- Το πρώτο βήμα είναι να σχηματίσουμε ένα "μεγάλο" μοντέλο με πολλές μεταβλητές στη δεξιά πλευρά
- Αυτό είναι γνωστό ως GUM (generalised unrestricted model/γενικευμένο «χωρίς περιορισμούς» μοντέλο)
- Σε αυτό το στάδιο, θέλουμε να βεβαιωθούμε ότι το μοντέλο πληροί όλες τις υποθέσεις/παραδοχές του γραμμικού υποδείγματος παλινδρόμησης
- Εάν παραβιαστούν οι υποθέσεις, πρέπει να λάβουμε τις κατάλληλες ενέργειες για να το διορθώσουμε, π.χ.
 - λήψη κορμών taking logs
 - προσθήκη χρονικών υστερήσεων
 - ψευδομεταβλητές (dummy variables)
- Πρέπει να το κάνουμε αυτό πριν δοκιμάσουμε υποθέσεις
- Μόλις έχουμε ένα μοντέλο που ικανοποιεί τις υποθέσεις, θα μπορούσε να είναι πολύ μεγάλο με πολλές χρονικές υστερήσεις και ανεξάρτητες μεταβλητές

Το επόμενο στάδιο είναι η επαναπαραμετροποίηση του μοντέλου

- Αποκλείοντας στατιστικά ασήμαντες ανεξάρτητες μεταβλητές
- ορισμένοι συντελεστές μπορεί να είναι στατιστικά ασήμαντα διαφορετικοί μεταξύ τους, ώστε να μπορούμε να τους συνδυάσουμε.
- Σε κάθε στάδιο, πρέπει να ελέγξουμε ότι οι υποθέσεις/παραδοχές είναι ακόμα εντάξει.
- Ας ελπίσουμε ότι σε αυτό το στάδιο, έχουμε ένα στατιστικά επαρκές εμπειρικό μοντέλο για το οποίο μπορούμε να χρησιμοποιήσουμε
 - ❖ έλεγχο υποκείμενων οικονομικών θεωριών
 - ❖ πρόβλεψη μελλοντικών τιμών της εξαρτημένης μεταβλητής

Εμπειρικό παράδειγμα II

Οι παράγοντες που καθορίζουν τις κρατικές αξιολογήσεις πιστοληπτικής ικανότητας

- Cantor and Packer (1996)
- Τι είναι οι κρατικές αξιολογήσεις πιστοληπτικής ικανότητας (sovereign credit ratings) και γιατί μας ενδιαφέρουν;
- Δύο οργανισμοί αξιολόγησης (Moody's and Standard and Poor's) παρέχουν πιστοληπτική αξιολόγηση για πολλές κυβερνήσεις.
- Κάθε πιθανή βαθμολογία συμβολίζεται με μια βαθμολογία:

Moody's
Aaa
.....
B3

Standard and Poor's
AAA
.....
B-

Στόχος της επιστημονικής εργασίας

- να επιχειρήσει να εξηγήσει και να μοντελοποιήσει τον τρόπο με τον οποίο κατέληξαν οι οργανισμοί αξιολόγησης τις βαθμολογίες τους.
- να χρησιμοποιούν τους ίδιους παράγοντες για να εξηγήσουν τις διαφορές των αποδόσεων των κρατικών αποδόσεων πάνω από ένα επιτόκιο μηδενικού κινδύνου.
- Να προσδιορίσει τους παράγοντες που επηρεάζουν τον τρόπο αντίδρασης των κρατικών αποδόσεων στις ανακοινώσεις αξιολόγησης.

Πιθανοί παράγοντες που καθορίζουν τις κρατικές αξιολογήσεις πιστοληπτικής ικανότητας

- Data

Quantifying the ratings (dependent variable): Aaa/AAA=16, ... , B3/B-=1

- Explanatory variables (units of measurement):

- Per capita income in 1994 (thousands of dollars)

- Average annual GDP growth 1991-1994 (%)

- Average annual inflation 1992-1994 (%)

- Fiscal balance: Average annual government budget surplus as a proportion of GDP 1992-1994 (%)

- External balance: Average annual current account surplus as a proportion of GDP 1992-1994 (%)

- External debt Foreign currency debt as a proportion of exports 1994 (%)

- Dummy for economic development

- Dummy for default history

Income and inflation are transformed to their logarithms.

Γραμμικό Υπόδειγμα Παλινδρόμησης - Εκτίμηση με τη μέθοδο ελαχίστων τετραγώνων

Explanatory Variable	Expected sign	Dependent Variable			
		Average Rating	Moody's Rating	S&P Rating	Moody's / S&P Difference
Intercept	?	1.442 (0.663)	3.408 (1.379)	-0.524 (-0.223)	3.932** (2.521)
Per capita income	+	1.242*** (5.302)	1.027*** (4.041)	1.458*** (6.048)	-0.431*** (-2.688)
GDP growth	+	0.151 (1.935)	0.130 (1.545)	0.171** (2.132)	-0.040 (0.756)
Inflation	-	-0.611*** (-2.839)	-0.630*** (-2.701)	-0.591*** (2.671)	-0.039 (-0.265)
Fiscal Balance	+	0.073 (1.324)	0.049 (0.818)	0.097* (1.71)	-0.048 (-1.274)
External Balance	+	0.003 (0.314)	0.006 (0.535)	0.001 (0.046)	0.006 (0.779)
External Debt	-	-0.013*** (-5.088)	-0.015*** (-5.365)	-0.011*** (-4.236)	-0.004*** (-2.133)
Development dummy	+	2.776*** (4.25)	2.957*** (4.175)	2.595*** (3.861)	0.362 (0.81)
Default dummy	-	-2.042*** (-3.175)	-1.63** (-2.097)	-2.622*** (-3.962)	1.159*** (2.632)
Adjusted R^2		0.924	0.905	0.926	0.836

Notes: t -ratios in parentheses; *, **, and *** indicate significance at the 10%, 5% and 1% levels respectively. Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

Ερμηνεία του Μοντέλου

Από στατιστική άποψη (statistical perspective)

- Μη διαγνωστικοί έλεγχοι.
- Ο προσαρμοσμένος R^2 είναι υψηλός.
- Κοιτάξτε τα κατάλοιπα : πραγματική βαθμολογία - προσαρμοσμένη βαθμολογία

Από χρηματοοικονομική άποψη (financial perspective)

- Έχουν οι συντελεστές τα αναμενόμενα πρόσημα και μεγέθη τους;

Προσθέτουν οι Βαθμολογίες στις Διαθέσιμες Δημόσιες Πληροφορίες;

Η εξαρτημένη μεταβλητή είναι

- Log (Yield on the sovereign bond - yield on a US treasury bond)

Προσθέτουν οι Βαθμολογίες στις Διαθέσιμες Δημόσιες Πληροφορίες; Αποτελέσματα

Variable	Expected Sign	Dependent Variable: Log (yield spread)		
		(1)	(2)	(3)
Intercept	?	2.105*** (16.148)	0.466 (0.345)	0.074 (0.071)
Average Rating	-	-0.221*** (-19.175)		-0.218*** (-4.276)
Per capita income	-		-0.144 (-0.927)	0.226 (1.523)
GDP growth	-		-0.004 (-0.142)	0.029 (1.227)
Inflation	+		0.108 (1.393)	-0.004 (-0.068)
Fiscal Balance	-		-0.037 (-1.557)	-0.02 (-1.045)
External Balance	-		-0.038 (-1.29)	-0.023 (-1.008)
External Debt	+		0.003*** (2.651)	0.000 (0.095)
Development dummy	-		-0.723*** (-2.059)	-0.38 (-1.341)
Default dummy	+		0.612*** (2.577)	0.085 (0.385)
Adjusted R^2		0.919	0.857	0.914

Notes: t-ratios in parentheses; *, **, and *** indicate significance at the 10%, 5% and 1% levels respectively. Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

Ποιοι παράγοντες καθορίζουν την αντίδραση της αγοράς στις ανακοινώσεις αξιολογήσεων;

- Το δείγμα: Κάθε ανακοίνωση αλλαγής βαθμολογίας που συνέβη μεταξύ 1987 και 1994: 79 τέτοιες ανακοινώσεις σε 18 χώρες.
- 39 ήταν πραγματικές αλλαγές αξιολογήσεων.
- 40 ήταν αλλαγές στη λίστα παρακολούθησης / προοπτικές.
- Η εξαρτημένη μεταβλητή: μεταβολές των spreads σε σχέση με το αμερικανικό ομόλογο σε διάστημα 2 ημερών κατά τη στιγμή της ανακοίνωσης.

Ποιοι παράγοντες καθορίζουν την αντίδραση της αγοράς στις ανακοινώσεις αξιολογήσεων; Ερμηνευτικές μεταβλητές

0 /1 ψευδομεταβλητών (dummies) για

- Αν η ανακοίνωση ήταν θετική
- Αν υπήρξε πραγματική αλλαγή βαθμολογίας.
- Αν το ομόλογο ήταν κερδοσκοπικού βαθμού
- Είχε υπήρξε άλλη ανακοίνωση αξιολόγησης τις προηγούμενες 60 ημέρες.

και

- Η αλλαγή στο spread κατά τις προηγούμενες 60 ημέρες.
- Το κενό αξιολόγησης (ratings gap) μεταξύ του του οργανισμού που ανακοίνωσε την αξιολόγηση και του άλλου οργανισμού

Ποιοι παράγοντες καθορίζουν την αντίδραση της αγοράς στις ανακοινώσεις αξιολογήσεων; Αποτελέσματα

Dependent Variable: Log Relative Spread	
Independent variable	Coefficient (<i>t</i> -ratio)
Intercept	-0.02 (-1.4)
Positive announcements	0.01 (0.34)
Ratings changes	-0.01 (-0.37)
Moody's announcements	0.02 (1.51)
Speculative grade	0.03** (2.33)
Change in relative spreads from day -60 to day -1	-0.06 (-1.1)
Rating gap	0.03* (1.7)
Other rating announcements from day -60 to day -1	0.05** (2.15)
Adjusted R^2	0.12

Note: * and ** denote significance at the 10% and 5% levels respectively. Source: Cantor and Packer (1996). Reprinted with permission from *Institutional Investor*.

Συμπεράσματα

- 6 παράγοντες φαίνεται να παίζουν μεγάλο ρόλο στον καθορισμό της πιστοληπτικής ικανότητας των κρατών - incomes, GDP growth, inflation, external debt, industrialised or not, and default history.
- Οι αξιολογήσεις παρέχουν περισσότερες πληροφορίες σχετικά με τις αποδόσεις από όλους τους μακροοικονομικούς παράγοντες μαζί.
- Δεν μπορούμε να καθορίσουμε καλά ποιοι παράγοντες επηρεάζουν τον τρόπο με τον οποίο οι αγορές θα αντιδράσουν στις ανακοινώσεις αξιολογήσεων.

Σχόλια πάνω στην επιστημονική εργασία

- Μόνο 49 παρατηρήσεις για το πρώτο σετ παλινδρομήσεων και 35 για τις παλινδρομήσεις της απόδοσης και μέχρι 10 ανεξάρτητες μεταβλητές.
- Καμία προσπάθεια επαναπαραμετροποίησης.
- Μικρή προσπάθεια διαγνωστικού ελέγχου.
- Από πού προέρχονται οι παράγοντες (ερμηνευτικές μεταβλητές);