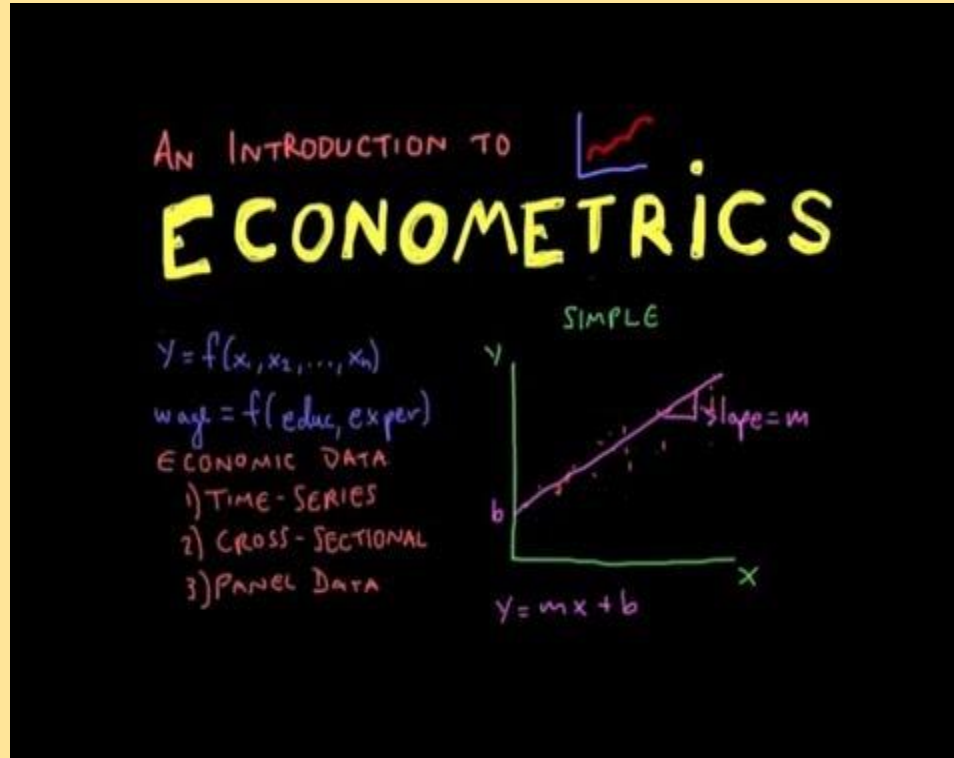


Οικονομικό Πανεπιστήμιο Αθηνών

Μάθημα: "Εφαρμογές Στατιστικών Μεθόδων σε Επιχειρησιακά Προβλήματα"

Συνάντηση: 13η



Μερικός Συντελεστής Προσδιορισμού

***Ποια είναι η συνεισφορά
της νέας (ή των νέων)
ανεξάρτητων μεταβλητών
στην ερμηνεία του ανερμήνευτου
κομμματιού της μεταβλητότητας της Ψ ;***

↗ μερικός συντελεστής προσδιορισμού

R (partial) ² =

$$= [R \text{ (full) } ^2 - R \text{ (reduced) } ^2] / [1 - R \text{ (reduced) } ^2]$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = -1,97 + 0,64 * \log(\text{αστικός_πληθυσμός})$
(k=1) **(R²=0,6169)**

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,22 + 0,67 * \log(\text{αστικός_πληθυσμός})$
(k=2) $- 1,95 * \log(\text{δείκτης_οικονομ_ανισοτ})$
(R²=0,6562)

↗ μερικός συντελεστής προσδιορισμού

η προσθήκη της νέας μεταβλητής
εξηγεί το 5,6% της ανερμήνευτης
μεταβλητότητας που δεν εξηγούσε
το αρχικό μοντέλο

R (partial) ^2 =

$$= [0,6562 - 0,6169] / [1 - 0,6169] = 0,0393 / 0,3831 = \mathbf{0,102 (10,2\%)}$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = -1,97 + 0,64 * \log(\text{αστικός_πληθυσμός})$
(**R^2=0,6169**)

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,22 + 0,67 * \log(\text{αστικός_πληθυσμός})$
- 1,95 * $\log(\text{δείκτης_οικονομ_ανισοτ})$
(**R^2=0,6562**)

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 13:52

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.979340	1.195865	-1.655154	0.1046
LOG(URBAN_POPULATION)	0.641322	0.073709	8.700733	0.0000
R-squared	0.616961	Mean dependent var	8.374907	
Adjusted R-squared	0.608811	S.D. dependent var	1.319081	
S.E. of regression	0.825021	Akaike info criterion	2.493143	
Sum squared resid	31.99097	Schwarz criterion	2.570360	
Log likelihood	-59.08201	Hannan-Quinn criter.	2.522439	
F-statistic	75.70276	Durbin-Watson stat	2.073934	
Prob(F-statistic)	0.000000			

Αρχικό Μοντέλο:Ένα (1-0,6169)
= 0,3831**= 38,31%***μένει ανερμήνευτο*

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 14:08

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.224932	2.939727	1.437185	0.1574
LOG(URBAN_POPULATION)	0.679244	0.072500	9.368902	0.0000
LOG(GINI_INDEX)	-1.955313	0.853282	-2.291519	0.0266

R-squared	0.656206	Mean dependent var	8.374907
Adjusted R-squared	0.641258	S.D. dependent var	1.319081
S.E. of regression	0.790064	Akaike info criterion	2.425865
Sum squared resid	28.71325	Schwarz criterion	2.541690
Log likelihood	-56.43368	Hannan-Quinn criter.	2.469809
F-statistic	43.90050	Durbin-Watson stat	2.160405
Prob(F-statistic)	0.000000		

Νέο Μοντέλο:

(0,6562 - 0,6169)

= 0,0393

(0,0393 / 0,3831)

= 10,2%

↗ μερικός συντελεστής προσδιορισμού

R (partial) ^2 =

$$= [R \text{ (full) } ^2 - R \text{ (reduced) } ^2] / [1 - R \text{ (reduced) } ^2]$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = -1,97 + 0,64 * \log(\text{αστικός_πληθυσμός})$
(k=1) **(R^2=0,6169)**

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,81 + 0,71 * \log(\text{αστικός_πληθυσμός})$
(k=3) $- 1,71 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,69 * \log(\text{ατμοσφ_ρύπανση})$
(R^2=0,7198)

→ μερικός συντελεστής προσδιορισμού

η προσθήκη των δύο νέων μεταβλητών εξηγεί το 26% της ανερμήνευτης μεταβλητότητας που δεν εξηγούσε το αρχικό μοντέλο

R (partial) ^2 =

$$= [0,7198 - 0,6169] / [1 - 0,6169] = 0,1029 / 0,3831 = \mathbf{0,26 \text{ (26\%)}}$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = -1,97 + 0,64 * \log(\text{αστικός_πληθυσμός})$
(R^2=0,6169)

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,81 + 0,71 * \log(\text{αστικός_πληθυσμός})$
 $- 1,71 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,69 * \log(\text{ατμοσφ_ρύπανση})$
(R^2=0,7198)

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 13:52

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.979340	1.195865	-1.655154	0.1046
LOG(URBAN_POPULATION)	0.641322	0.073709	8.700733	0.0000
R-squared	0.616961	Mean dependent var	8.374907	
Adjusted R-squared	0.608811	S.D. dependent var	1.319081	
S.E. of regression	0.825021	Akaike info criterion	2.493143	
Sum squared resid	31.99097	Schwarz criterion	2.570360	
Log likelihood	-59.08201	Hannan-Quinn criter.	2.522439	
F-statistic	75.70276	Durbin-Watson stat	2.073934	
Prob(F-statistic)	0.000000			

Αρχικό Μοντέλο:

Ένα (1-0,6169)
= 0,3831

= 38,31%

μένει ανερμήνευτο

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 14:16

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.814663	2.689311	1.790296	0.0801
LOG(URBAN_POPULATION)	0.719023	0.067327	10.67953	0.0000
LOG(GINI_INDEX)	-1.717268	0.782309	-2.195126	0.0334
LOG(PM2_5_AIR_POLLUTION__MEAN_A...	-0.692491	0.216561	-3.197678	0.0025
R-squared	0.719861	Mean dependent var	8.374907	
Adjusted R-squared	0.701185	S.D. dependent var	1.319081	
S.E. of regression	0.721062	Akaike info criterion	2.261925	
Sum squared resid	23.39688	Schwarz criterion	2.416359	
Log likelihood	-51.41716	Hannan-Quinn criter.	2.320517	
F-statistic	38.54478	Durbin-Watson stat	2.221986	
Prob(F-statistic)	0.000000			

Νέο Μοντέλο:

(0,7198 - 0,6169)

= 0,1029

(0,1029 / 0,3831)

= 26%

→ μερικός συντελεστής προσδιορισμού

R (partial) ² =

$$= [R \text{ (full) } ^2 - R \text{ (reduced) } ^2] / [1 - R \text{ (reduced) } ^2]$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = 4,22 + 0,67 * \log(\text{αστικός_πληθυσμός})$
(k=2)
 $- 1,95 * \log(\text{δείκτης_οικονομ_ανισοτ})$
(R²=0,6562)

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,81 + 0,71 * \log(\text{αστικός_πληθυσμός})$
(k=3)
 $- 1,71 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,69 * \log(\text{ατμοσφ_ρύπανση})$
(R²=0,7198)

→ μερικός συντελεστής προσδιορισμού

R (partial) ^2 =

$$= [0,7198 - 0,6562] / [1 - 0,6562] = 0,0636 / 0,3438 = \mathbf{0,184 (18,4\%)}$$

↙
νέος

↙
παλιός

↙
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = 4,22 + 0,67 * \log(\text{αστικός_πληθυσμός})$
 $- 1,95 * \log(\text{δείκτης_οικονομ_ανισοτ})$
(R^2=0,6562)

full: $\log(\text{έκταση_σιδηροδρόμου}) = 4,81 + 0,71 * \log(\text{αστικός_πληθυσμός})$
 $- 1,71 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,69 * \log(\text{ατμοσφ_ρύπανση})$
(R^2=0,7198)

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 14:26

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.224932	2.939727	1.437185	0.1574
LOG(URBAN_POPULATION)	0.679244	0.072500	9.368902	0.0000
LOG(GINI_INDEX)	-1.955313	0.853282	-2.291519	0.0266
R-squared	0.656206	Mean dependent var	8.374907	
Adjusted R-squared	0.641258	S.D. dependent var	1.319081	
S.E. of regression	0.790064	Akaike info criterion	2.425865	
Sum squared resid	28.71325	Schwarz criterion	2.541690	
Log likelihood	-56.43368	Hannan-Quinn criter.	2.469809	
F-statistic	43.90050	Durbin-Watson stat	2.160405	
Prob(F-statistic)	0.000000			

Αρχικό Μοντέλο:

$$\begin{aligned} \text{Ένα } (1-0,6562) \\ &= 0,3438 \\ &= 34,38\% \end{aligned}$$

μένει ανερμήνευτο

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 14:16

Sample: 1 49

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.814663	2.689311	1.790296	0.0801
LOG(URBAN_POPULATION)	0.719023	0.067327	10.67953	0.0000
LOG(GINI_INDEX)	-1.717268	0.782309	-2.195126	0.0334
LOG(PM2_5_AIR_POLLUTION__MEAN_A...	-0.692491	0.216561	-3.197678	0.0025
R-squared	0.719861	Mean dependent var	8.374907	
Adjusted R-squared	0.701185	S.D. dependent var	1.319081	
S.E. of regression	0.721062	Akaike info criterion	2.261925	
Sum squared resid	23.39688	Schwarz criterion	2.416359	
Log likelihood	-51.41716	Hannan-Quinn criter.	2.320517	
F-statistic	38.54478	Durbin-Watson stat	2.221986	
Prob(F-statistic)	0.000000			

Νέο Μοντέλο:

(0,7198 - 0,6562)

= 0,0636

(0,0636 / 0,3438)

= 18,49%

→ μερικός συντελεστής προσδιορισμού

R (partial) ^2 =

$$= [R \text{ (full)} ^2 - R \text{ (reduced)} ^2] / [1 - R \text{ (reduced)} ^2]$$

↓
νέος

↓
παλιός

↓
παλιός

reduced: $\log(\text{έκταση_σιδηροδρόμου}) = 4,81 + 0,71 * \log(\text{αστικός_πληθυσμός})$
 (k=3)
 $- 1,71 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,69 * \log(\text{ατμοσφ_ρύπανση})$
 (διορθωμένο **R^2=0,70**)

full: $\log(\text{έκταση_σιδηροδρόμου}) = 6,02 + 0,69 * \log(\text{αστικός_πληθυσμός})$
 (k=4)
 $- 1,86 * \log(\text{δείκτης_οικονομ_ανισοτ})$
 $- 0,86 * \log(\text{ατμοσφ_ρύπανση})$
 $+ 0,04 * \log(\text{προσφ_εξωτερ})$
 (διορθωμένο **R^2=0,68**)

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 14:34

Sample: 1 49

Included observations: 48

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	6.028217	2.842018	2.121104	0.0397
LOG(URBAN_POPULATION)	0.690555	0.069547	9.929306	0.0000
LOG(GINI_INDEX)	-1.864342	0.797472	-2.337816	0.0241
LOG(PM2_5_AIR_POLLUTION__MEAN_A...	-0.864983	0.284448	-3.040919	0.0040
LOG(REFUGEE_POPULATION_BY_COU...	0.043798	0.052839	0.828897	0.4117
R-squared	0.707969	Mean dependent var	8.432368	
Adjusted R-squared	0.680803	S.D. dependent var	1.269553	
S.E. of regression	0.717266	Akaike info criterion	2.271591	
Sum squared resid	22.12220	Schwarz criterion	2.466508	
Log likelihood	-49.51819	Hannan-Quinn criter.	2.345251	
F-statistic	26.06116	Durbin-Watson stat	2.215620	
Prob(F-statistic)	0.000000			

Υποθέσεις του Υποδείγματος

Υπόθεση 1η: Κανονικότητα

Μια από τις υποθέσεις του υποδείγματος της γραμμικής παλινδρόμησης είναι ότι ο **διαταρακτικός όρος κατανέμεται κανονικά** με μέσο μηδέν και σταθερή διακύμανση.

Συνέπειες Παραβίασης: Οι συντελεστές της παλινδρόμησης συνήθως **ΔΕΝ** επηρεάζονται. Επηρεάζονται όμως τα **τυπικά σφάλματα** των εκτιμητών. ***Αυτό σημαίνει πως οι έλεγχοι για την στατιστική σημαντικότητα και τα διαστήματα εμπιστοσύνης μπορεί να οδηγούν σε λαθεμένα συμπεράσματα . . .***



Carlos M. Jarque Uribe



*Έλεγχος για την
παραβίαση (ή μη)
της υπόθεσης της
κανονικότητας ...*

Anil K. Bera



Bera in 2024

Οι **Jarque και Bera (1987)** πρότειναν μια μέθοδο που βασίζεται στα **κατάλοιπα** που προκύπτουν από τη μέθοδο των ελαχίστων τετραγώνων.

Ο έλεγχος αυτός ακολουθεί τα παρακάτω βήματα:

- Εκτιμούμε το υπόδειγμα με τη μέθοδο των ελαχίστων τετραγώνων και σώζουμε τα κατάλοιπα e_t .
- Χρησιμοποιώντας τα κατάλοιπα υπολογίζουμε τους συντελεστές της ασυμμετρίας και της κύρτωσης.

Αν τα **κατάλοιπα** ακολουθούν την κανονική κατανομή τότε θα έχω **$S = 0$** και **$K = 3$** .

Έλεγχος της Κανονικότητας του Διαταρακτικού Όρου

Οι δύο υποθέσεις μηδενική (*null hypothesis*) και εναλλακτική) (*alternative hypothesis*) του ελέγχου μπορούν να διατυπωθούν ως εξής:

- H_0 : Τα κατάλοιπα ακολουθούν την κανονική κατανομή (κατανέμονται κανονικά).
- H_a : Τα κατάλοιπα δεν ακολουθούν την κανονική κατανομή (δεν κατανέμονται κανονικά).

Έλεγχος της Κανονικότητας του Διαταρακτικού Όρου

Υπολογίζουμε το στατιστικό των **Jarque– Bera**
ως εξής:

$$JB = \frac{n}{6} \left[S^2 + \frac{1}{4} (K - 3)^2 \right]$$

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/09/25 Time: 20:54

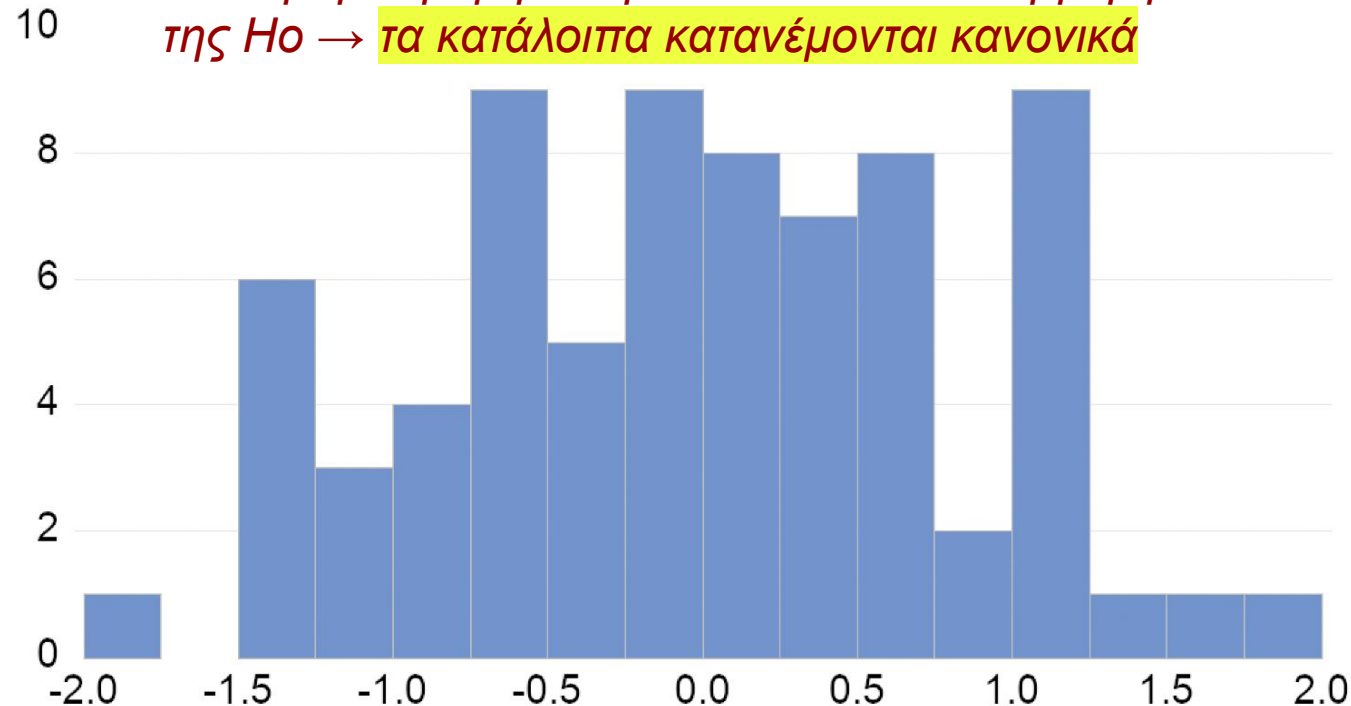
Sample (adjusted): 11 258

Included observations: 74 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.961210	-2.643077	0.0101
LOG(URBAN_POPULATION)	0.665739	0.059392	11.20932	0.0000
R-squared	0.635718	Mean dependent var	8.180012	
Adjusted R-squared	0.630658	S.D. dependent var	1.359800	
S.E. of regression	0.826398	Akaike info criterion	2.483175	
Sum squared resid	49.17122	Schwarz criterion	2.545447	
Log likelihood	-89.87746	Hannan-Quinn criter.	2.508016	
F-statistic	125.6489	Durbin-Watson stat	1.639665	
Prob(F-statistic)	0.000000			

Έλεγχος Κανονικότητας στο enviews

Πολύ μεγάλη τιμή του p -value $\rightarrow$ ΜΗ απόρριψη της $H_0 \rightarrow$ τα κατάλοιπα κατανέμονται κανονικά



Series: Residuals

Sample 11-258

Observations 74

Mean -3.03e-16

Median 0.014540

Maximum 1.845413

Minimum -1.771510

Std. Dev. 0.820718

Skewness -0.042831

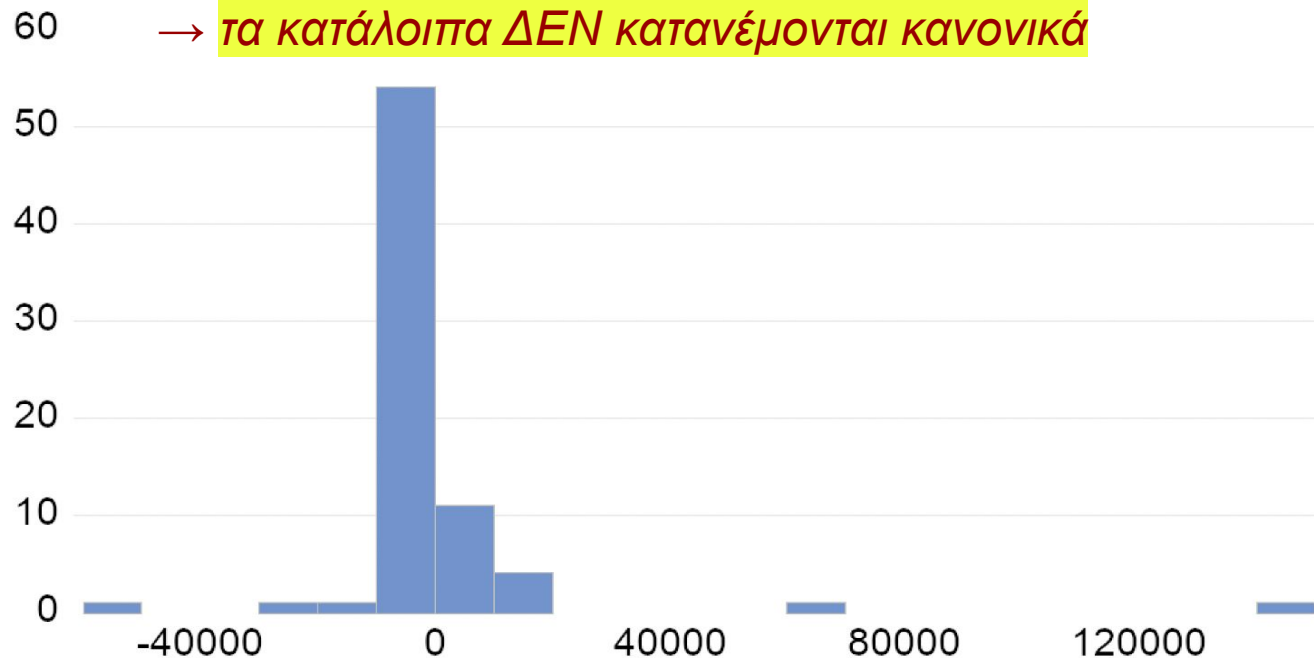
Kurtosis 2.300971

Jarque-Bera 1.529268

Probability 0.465504

Έλεγχος Κανονικότητας στο *enviews* (μοντέλο ΧΩΡΙΣ λογαρίθμηση)

Πολύ μικρή τιμή του *p-value* → απόρριψη της H_0
→ τα κατάλοιπα ΔΕΝ κατανέμονται κανονικά



Series: Residuals

Sample 11 258

Observations 74

Mean -8.23e-13

Median -3265.074

Maximum 146181.3

Minimum -54080.81

Std. Dev. 20598.73

Skewness 5.061570

Kurtosis 37.09950

Jarque-Bera 3901.200

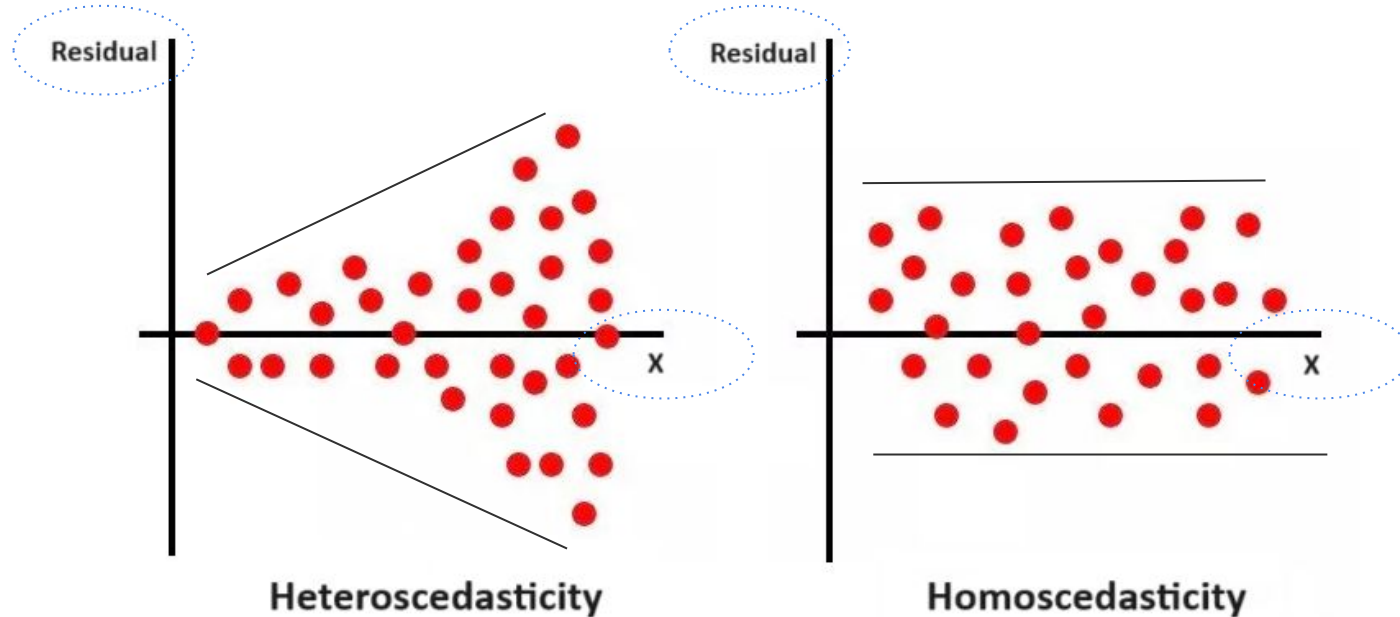
Probability 0.000000

Υπόθεση 2η: Ομοσκεδαστικότητα

Μια άλλη βασική υπόθεση της γραμμικής παλινδρόμησης είναι ότι η διακύμανση του διαταρακτικού όρου **είναι παραμένει σταθερή**

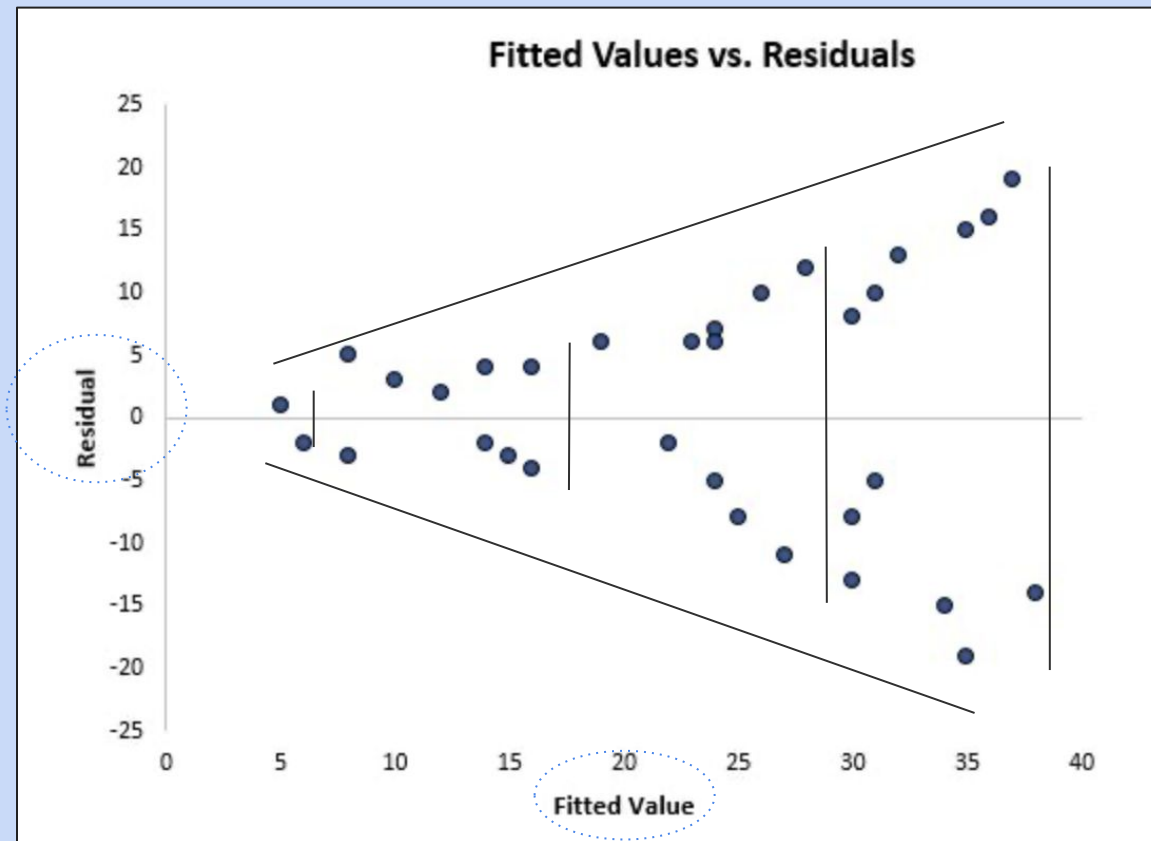
Συνέπειες Παραβίασης: Τα **τυπικά σφάλματα** των εκτιμητών μπορεί να είναι ανακριβή. **Ειδικότερα, η ετεροσκεδαστικότητα μπορεί να οδηγήσει σε υποεκτίμηση ή υπερεκτίμηση των τυπικών σφαλμάτων των εκτιμητών.**
Αυτό σημαίνει πως οι έλεγχοι για την στατιστική σημαντικότητα και τα διαστήματα εμπιστοσύνης μπορεί να οδηγούν σε λαθεμένα συμπεράσματα . . .

Διαγραμματικές Ενδείξεις

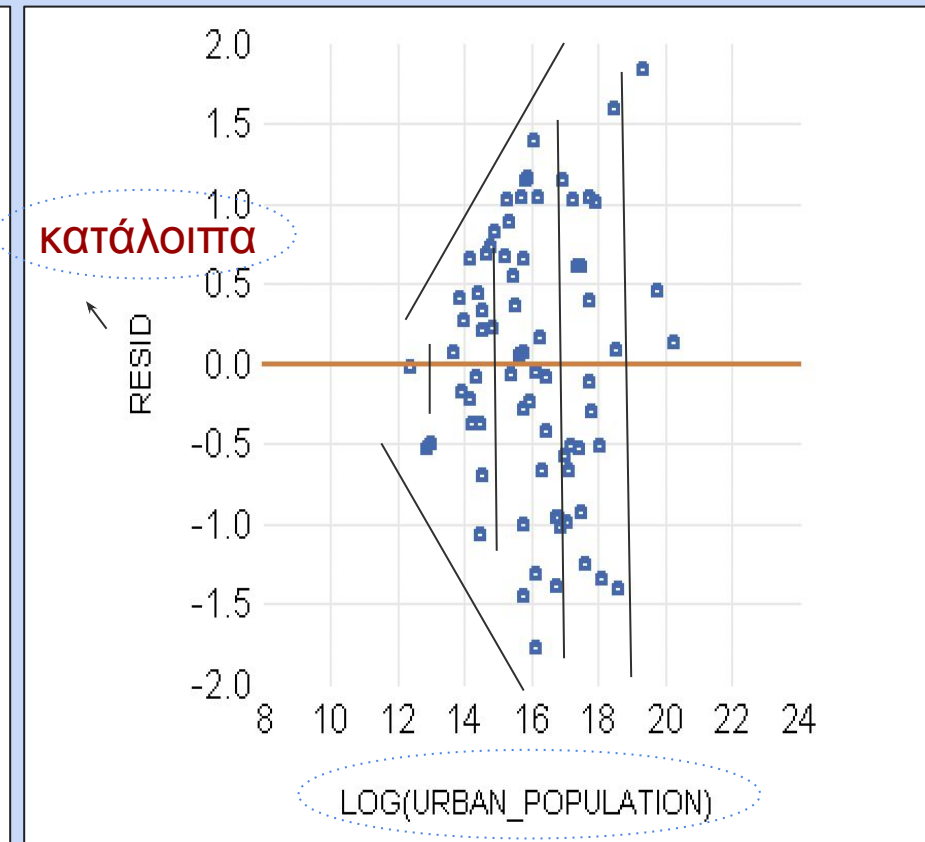
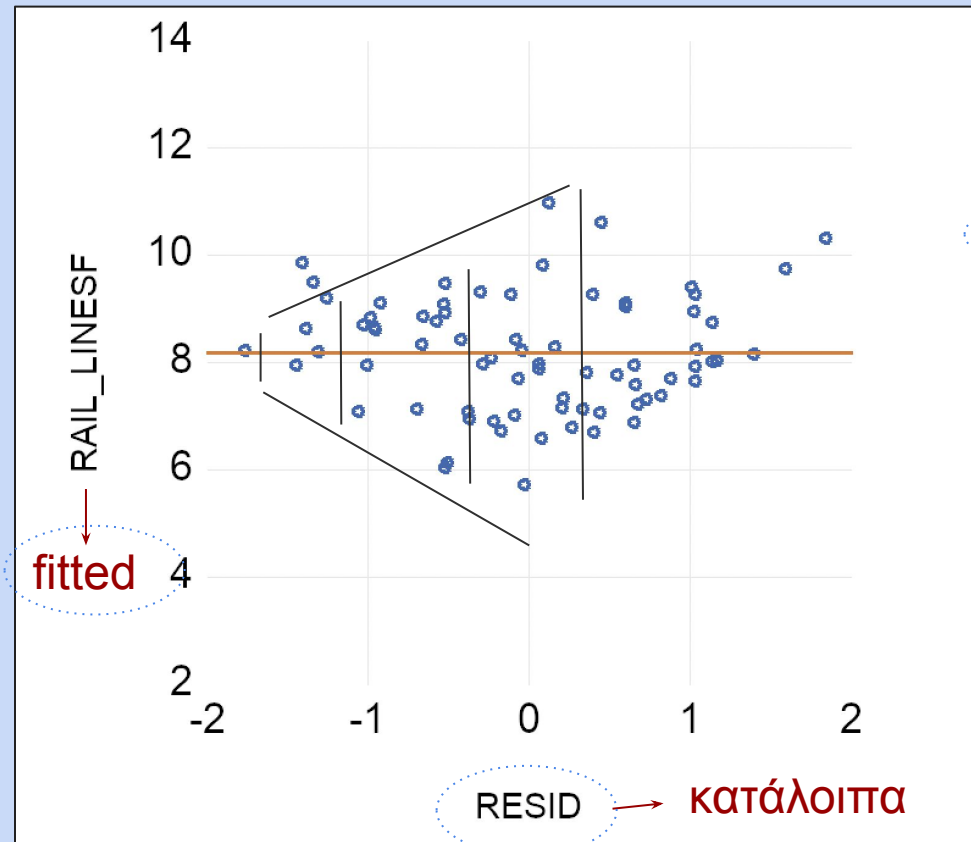


Βάλτε στο διάγραμμα διασποράς τις τιμές της ανεξάρτητης μεταβλητής (X_i) και τις τιμές των καταλοίπων (u_i)

Διαγραμματικές Ενδείξεις



Διαγραμματικές Ενδείξεις



ΕΛΕΓΧΟΣ BREUSCH-PAGAN-GODFREY

Αν υπάρχουν **ερμηνευτικές μεταβλητές** που επηρεάζουν τη **διακύμανση του διαταρακτικού όρου**, τότε η διακύμανση είναι μια συνάρτηση των ερμηνευτικών μεταβλητών, δηλαδή έχουμε:

$$\sigma_i^2 = a_0 + a_1 X_{1i} + a_2 X_{2i} + \dots + a_k X_{ki} + v_t$$

ΕΛΕΓΧΟΣ BREUSCH-PAGAN-GODFREY

Διαδικασία ελέγχου:

1η Φάση: με την MET, εκτιμούμε το αρχικό υπόδειγμα: $Y = X\beta + \varepsilon$

2η Φάση: τυποποιούμε τα κατάλοιπα με τον εκτιμητή της διακύμανσης
του ε

3η Φάση: με την MET, εκτιμούμε το παρακάτω μοντέλο:

$$\hat{u}_i^2 = a_0 + a_1 X_{1i} + a_2 X_{2i} + \dots + a_k X_{ki} + v_t$$

ΕΛΕΓΧΟΣ BREUSCH-PAGAN-GODFREY

Διαδικασία ελέγχου:

4η Φάση: Υπολογίζουμε την στατιστική $LM = nR^2$
όπου R^2 προκύπτει από το 2ο μοντέλο.

5η Φάση: Αν $LM > X^2$ ($\alpha\%$, k) όπου k = αριθμός μεταβλητών του μοντέλου [2], τότε δεχόμαστε ότι υπάρχει γραμμική ετεροσκεδαστικότητα.

ΕΛΕΓΧΟΣ BREUSCH-PAGAN-GODFREY

Heteroskedasticity Test: Breusch-Pagan-Godfrey

Null hypothesis: Homoskedasticity

ύπαρξη ετεροσκεδαστικότητας

F-statistic	9.605412	Prob. F(1,72)	0.0028
Obs*R-squared	8.710213	Prob. Chi-Square(1)	0.0032
Scaled explained SS	5.363744	Prob. Chi-Square(1)	0.0206

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 04/09/25 Time: 22:06

Sample: 11 258

Included observations: 74

→ τα τετράγωνα των καταλοίπων

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-1.924193	0.839457	-2.292188	0.0248
LOG(URBAN_POPULATION)	0.160754	0.051869	3.099260	0.0028
R-squared	0.117706	Mean dependent var	0.664476	
Adjusted R-squared	0.105451	S.D. dependent var	0.763076	
S.E. of regression	0.721721	Akaike info criterion	2.212300	
Sum squared resid	37.50347	Schwarz criterion	2.274572	
Log likelihood	-79.85509	Hannan-Quinn criter.	2.237141	
F-statistic	9.605412	Durbin-Watson stat	3.221013	
Prob(F-statistic)	0.002767			

ανεξάρτητη →

→ στατιστικά
σημαντικό

Μοντέλο για $k = 3$

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 15:08

Sample (adjusted): 11 258

Included observations: 49 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.814663	2.689311	1.790296	0.0801
LOG(URBAN_POPULATION)	0.719023	0.067327	10.67953	0.0000
LOG(GINI_INDEX)	-1.717268	0.782309	-2.195126	0.0334
LOG(PM2_5_AIR_POLLUTION__MEAN_A...	-0.692491	0.216561	-3.197678	0.0025
R-squared	0.719861	Mean dependent var	8.374907	
Adjusted R-squared	0.701185	S.D. dependent var	1.319081	
S.E. of regression	0.721062	Akaike info criterion	2.261925	
Sum squared resid	23.39688	Schwarz criterion	2.416359	
Log likelihood	-51.41716	Hannan-Quinn criter.	2.320517	
F-statistic	38.54478	Durbin-Watson stat	2.380873	
Prob(F-statistic)	0.000000			

ΕΛΕΓΧΟΣ BREUSCH-PAGAN-GODFREY

Heteroskedasticity Test: Breusch-Pagan-Godfrey

Null hypothesis: Homoskedasticity

F-statistic	2.318207	Prob. F(3,45)	0.0882
Obs*R-squared	6.559118	Prob. Chi-Square(3)	0.0874
Scaled explained SS	3.671332	Prob. Chi-Square(3)	0.2992

Test Equation:

Dependent Variable: RESID^2

Method: Least Squares

Date: 04/10/25 Time: 15:09

Sample: 11 258

Included observations: 49

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-3.519274	1.992525	-1.766238	0.0841
LOG(URBAN_POPULATION)	0.099338	0.049883	1.991406	0.0525
LOG(GINI_INDEX)	0.727479	0.579617	1.255102	0.2159
LOG(PM2_5_AIR_POLLUTION__MEAN_A...	-0.048086	0.160451	-0.299693	0.7658

R-squared	0.133860	Mean dependent var	0.477487
Adjusted R-squared	0.076117	S.D. dependent var	0.555811
S.E. of regression	0.534239	Akaike info criterion	1.662161
Sum squared resid	12.84351	Schwarz criterion	1.816595
Log likelihood	-36.72294	Hannan-Quinn criter.	1.720753
F-statistic	2.318207	Durbin-Watson stat	1.769919
Prob(F-statistic)	0.088207		

Διόρθωση

Ετεροσκεδαστικότητας ($k = 1$)

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

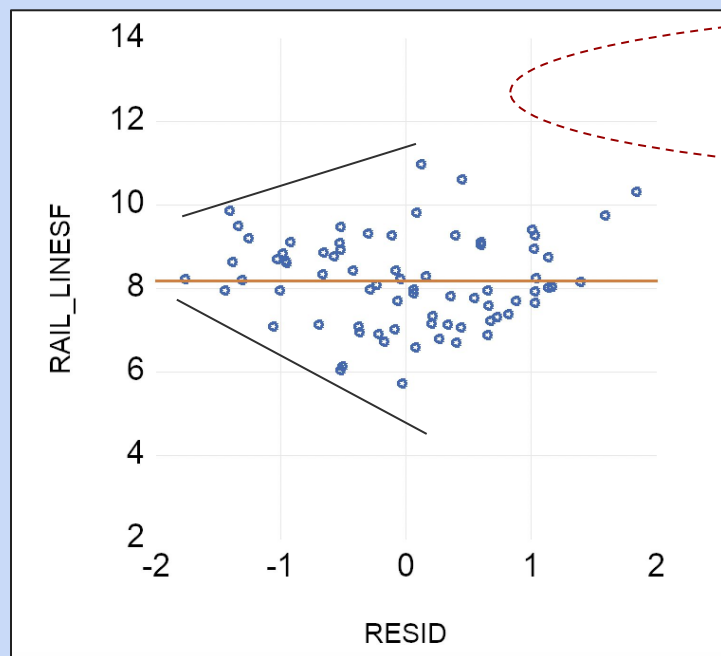
Date: 04/10/25 Time: 15:30

Sample (adjusted): 11 258

Included observations: 74 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.961210	-2.643077	0.0101
LOG(URBAN_POPULATION)	0.665739	0.059392	11.20932	0.0000
R-squared	0.635718	Mean dependent var		8.180012
Adjusted R-squared	0.630658	S.D. dependent var		1.359800
S.E. of regression	0.826398	Akaike info criterion		2.483175
Sum squared resid	49.17122	Schwarz criterion		2.545447
Log likelihood	-89.87746	Hannan-Quinn criter.		2.508016
F-statistic	125.6489	Durbin-Watson stat		1.639665
Prob(F-statistic)	0.000000			

Διόρθωση Ετεροσκεδαστικότητας



Heteroskedasticity Test: White
Null hypothesis: Homoskedasticity

F-statistic	5.306723	Prob. F(2,71)	0.0071
Obs*R-squared	9.623352	Prob. Chi-Square(2)	0.0081
Scaled explained SS	5.926055	Prob. Chi-Square(2)	0.0517

Test Equation:
Dependent Variable: $RESID^2$
Method: Least Squares
Date: 04/10/25 Time: 15:31
Sample: 11 258
Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-8.312904	6.421293	-1.294584	0.1997
$\text{LOG(URBAN_POPULATION)}^2$	-0.024520	0.024434	-1.003538	0.3190
$\text{LOG(URBAN_POPULATION)}$	0.956328	0.794464	1.203740	0.2327

R-squared	0.130045	Mean dependent var	0.664476
Adjusted R-squared	0.105540	S.D. dependent var	0.763076
S.E. of regression	0.721686	Akaike info criterion	2.225242
Sum squared resid	36.97895	Schwarz criterion	2.318650
Log likelihood	-79.33395	Hannan-Quinn criter.	2.262504
F-statistic	5.306723	Durbin-Watson stat	3.089265
Prob(F-statistic)	0.007114		

χωρίς
διόρθωσηΔιόρθωση
Ετεροσκεδαστικότηταςμε
διόρθωση

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)				
Method: Least Squares				
Date: 04/10/25 Time: 15:30				
Sample (adjusted): 11 258				
Included observations: 74 after adjustments				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.961210	-2.643077	0.0101
LOG(URBAN_POPULATION)	0.665739	0.059392	11.20932	0.0000
R-squared	0.635718	Mean dependent var		8.180012
Adjusted R-squared	0.630658	S.D. dependent var		1.359800
S.E. of regression	0.826398	Akaike info criterion		2.483175
Sum squared resid	49.17122	Schwarz criterion		2.545447
Log likelihood	-89.87746	Hannan-Quinn criter.		2.508016
F-statistic	125.6489	Durbin-Watson stat		1.639665
Prob(F-statistic)	0.000000			

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)				
Method: Least Squares				
Date: 04/10/25 Time: 15:35				
Sample (adjusted): 11 258				
Included observations: 74 after adjustments				
Huber-White-Hinkley (HC1) heteroskedasticity consistent standard errors and covariance				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.867658	-2.928054	0.0046
LOG(URBAN_POPULATION)	0.665739	0.056079	11.87143	0.0000
R-squared	0.635718	Mean dependent var		8.180012
Adjusted R-squared	0.630658	S.D. dependent var		1.359800
S.E. of regression	0.826398	Akaike info criterion		2.483175
Sum squared resid	49.17122	Schwarz criterion		2.545447
Log likelihood	-89.87746	Hannan-Quinn criter.		2.508016
F-statistic	125.6489	Durbin-Watson stat		1.639665
Prob(F-statistic)	0.000000	Wald F-statistic		140.9308
Prob(Wald F-statistic)	0.000000			

Διόρθωση Ετεροσκεδαστικότητας

με
διόρθωση

χωρίς
διόρθωση

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 15:30

Sample (adjusted): 11 258

Included observations: 74 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.961210	-2.643077	0.0101
LOG(URBAN_POPULATION)	0.665739	0.059392	11.20932	0.0000
R-squared	0.635718	Mean dependent var		8.180012
Adjusted R-squared	0.630658	S.D. dependent var		1.359800
S.E. of regression	0.826398	Akaike info criterion		2.483175
Sum squared resid	49.17122	Schwarz criterion		2.545447
Log likelihood	-89.87746	Hannan-Quinn criter.		2.508016
F-statistic	125.6489	Durbin-Watson stat		1.639665
Prob(F-statistic)	0.000000			

Dependent Variable: LOG(RAIL_LINES__TOTAL_ROUTE_KM_)

Method: Least Squares

Date: 04/10/25 Time: 15:39

Sample (adjusted): 11 258

Included observations: 74 after adjustments

HAC standard errors & covariance (Bartlett kernel, Newey-West fixed bandwidth = 4.0000)

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.540551	0.825817	-3.076408	0.0030
LOG(URBAN_POPULATION)	0.665739	0.054167	12.29040	0.0000
R-squared	0.635718	Mean dependent var		8.180012
Adjusted R-squared	0.630658	S.D. dependent var		1.359800
S.E. of regression	0.826398	Akaike info criterion		2.483175
Sum squared resid	49.17122	Schwarz criterion		2.545447
Log likelihood	-89.87746	Hannan-Quinn criter.		2.508016
F-statistic	125.6489	Durbin-Watson stat		1.639665
Prob(F-statistic)	0.000000	Wald F-statistic		151.0540
Prob(Wald F-statistic)	0.000000			