

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ



ATHENS UNIVERSITY
OF ECONOMICS
AND BUSINESS

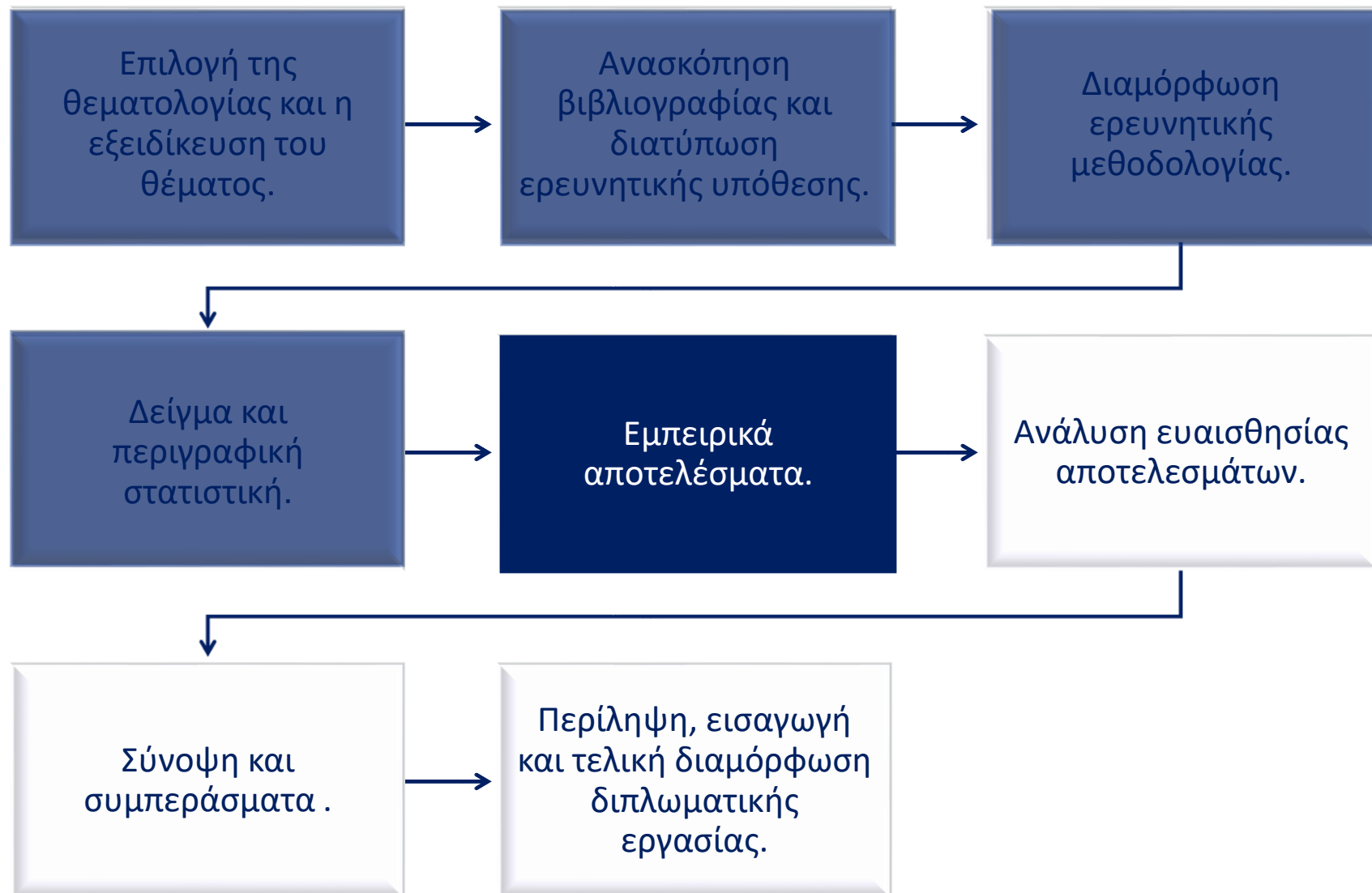
ΣΧΟΛΗ
ΔΙΟΙΚΗΣΗΣ
ΕΠΙΧΕΙΡΗΣΕΩΝ
SCHOOL OF
BUSINESS

ΜΕΤΑΠΤΥΧΙΑΚΟ ΛΟΓΙΣΤΙΚΗΣ &
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ
MSc IN ACCOUNTING & FINANCE

Σεμινάρια Εκπόνησης Διπλωματικής Εργασίας: Μέρος Γ



Εμπειρικά Αποτελέσματα



Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης

- Όπως είδαμε στο στάδιο της ερευνητικής μεθοδολογίας, η θεωρία όπως αυτή εκφράζεται με τη διατύπωση των ερευνητικών υποθέσεων λαμβάνει τη μορφή μίας μαθηματικής έκφρασης.
- Η περιγραφική στατιστική και η ανάλυση διακύμανσης δίδουν κάποιες κατευθύνσεις για την εγκυρότητα των ερευνητικών υποθέσεων.
- Ωστόσο το κυρίαρχο οικονομετρικό εργαλείο για τον έλεγχο των ερευνητικών υποθέσεων είναι η ανάλυση παλινδρόμησης.
- **Βασική ιδέα:** Έχοντας συλλέξει ένα δείγμα παρατηρήσεων είναι δυνατόν οι μαθηματικές εκφράσεις της θεωρίας να επιβεβαιωθούν για τον πληθυσμό;
- Η κλασσική γραμμική παλινδρόμηση είναι μία απόπειρα να απαντηθεί το παραπάνω ερώτημα διενεργώντας εκτίμηση της μαθηματικής έκφρασης της θεωρίας μας με την μέθοδο των ελαχίστων τετραγώνων.
 - Σημειώνεται ότι ο κλάδος της οικονομετρίας έχει επιδείξει ένα τεράστιο πλούτο τεχνικών εκτίμησης η παρουσίαση των οποίων αφενός ξεφεύγει από τους σκοπούς του σεμιναρίου και αφετέρου το κλασσικό γραμμικό μοντέλο παλινδρόμησης αποτελεί το εναρκτήριο βήμα οποιοδήποτε ανάλυσης.

Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης

- **Μονομεταβλητό υπόδειγμα γραμμικής παλινδρόμησης:** η εξαρτημένη μεταβλητή Y_i είναι συνάρτηση μίας μόνο ανεξάρτητης μεταβλητής (X_i):

$$Y_i = b_0 + b_1 X_i + u_i$$

- **Πολυμεταβλητό υπόδειγμα γραμμικής παλινδρόμησης:** η εξαρτημένη μεταβλητή Y_i είναι συνάρτηση περισσότερων από μίας ανεξάρτητων μεταβλητών (X_i):

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + \dots + b_n X_{ni} + u_i$$

Όπου:

Y_i : είναι οι τυχαίες τιμές της Y .

b_0 : είναι ο **σταθερός όρος** και αντιπροσωπεύει την τιμή της εξαρτημένης μεταβλητής Y_i όταν οι ανεξάρτητες μεταβλητές λάβουν την τιμή μηδέν $X_{1i} = X_{2i} = \dots = X_{ni} = 0$.

b_1, b_2, \dots, b_n : είναι οι **συντελεστές παλινδρόμησης**, οι τιμές των οποίων εκφράζουν την οριακή μεταβολή της Y όταν η αντίστοιχη ανεξάρτητη μεταβλητή μεταβληθεί κατά μία μονάδα (η ερμηνεία αυτή θέλει προσοχή ανάλογα με την εξειδίκευση του μοντέλου).

Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης

Όπου (συνέχεια):

u_i : είναι ο **διαταρακτικός όρος** και περιλαμβάνει:

- Ερμηνευτικές μεταβλητές που επηρεάζουν την εξαρτημένη μεταβλητή, αλλά δεν συμπεριελήφθησαν στο υπόδειγμα. Αιτίες παράληψης των μεταβλητών είναι (α) η έλλειψη διαθέσιμων τιμών, (β) η δυσκολία στην μέτρηση ορισμένων μεταβλητών και (γ) η “οικονομία” του μοντέλου
- Σφάλματα μέτρησης των υπό εξέταση μεταβλητών (εξαρτημένης ή ανεξάρτητης).
- Σφάλματα που προέρχονται από:
 - Λανθασμένη διατύπωση του υποδείγματος παλινδρόμησης,
 - Εσφαλμένο αρχικό θεωρητικό υπόδειγμα που έχει υιοθετηθεί για να περιγράψει το εν λόγω φαινόμενο (π.χ., μπορεί να έχει παραληφθεί από το μοντέλο η αλληλεπίδραση δυο μεταβλητών).

Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης

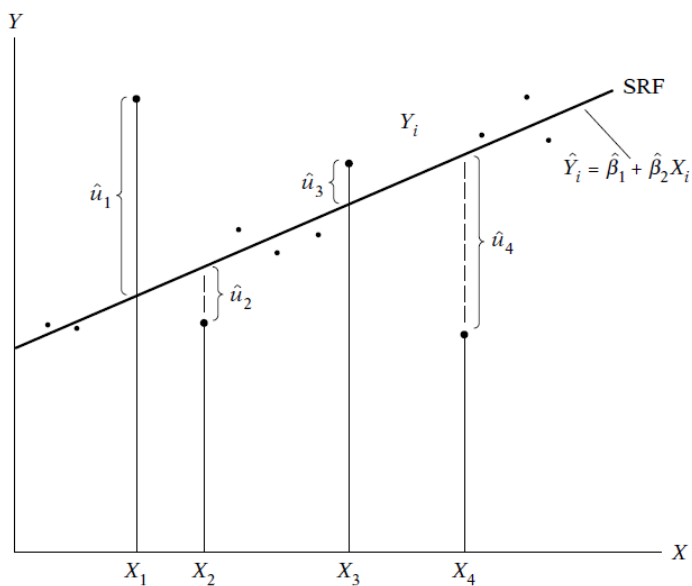
- Η εκτίμηση ενός γραμμικού μοντέλου παλινδρόμησης στοχεύει στην εκτίμηση των τιμών του σταθερού όρου (b_0) και των συντελεστών παλινδρόμησης (b_1, b_2, \dots, b_n), οι οποίοι ελαχιστοποιούν το άθροισμα των τετραγώνων των διαταρακτικών τιμών (OLS – Ordinary Least Squares Method).

Αρχική εξειδίκευση γραμμικού μοντέλου παλινδρόμησης	Εκτιμηθέν γραμμικό μοντέλο παλινδρόμησης
$Y_i = b_0 + b_1 X_i + e_i$	$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$
$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + \dots + b_n X_{ni} + e_i$	$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i} + \dots + \hat{b}_n X_{ni}$

- Οι εκτιμηθείσες τιμές του σταθερού όρου και των συντελεστών παλινδρόμησης αποδίδονται με τους όρους ($\hat{b}_0, \hat{b}_1, \hat{b}_2, \dots, \hat{b}_n$) αντίστοιχα.
- Για κάθε πραγματική τιμή της εξερχόμενης μεταβλητής (Y_i) διαμορφώνεται μία εκτιμηθείσα τιμή της ανεξάρτητης (\hat{Y}_i).

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης

Αρχική εξειδίκευση γραμμικού μοντέλου παλινδρόμησης	Εκτιμηθέν γραμμικό μοντέλο παλινδρόμησης
$Y_i = b_0 + b_1 X_i + u_i$	$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_i$
$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \dots + \dots + b_n X_{ni} + u_i$	$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{1i} + \hat{b}_2 X_{2i} + \dots + \hat{b}_n X_{ni}$



Εκτιμηθέν γραμμικό μοντέλο παλινδρόμησης
$\hat{u}_i = Y_i - \hat{Y}_i = Y_i - \hat{b}_0 + \hat{b}_1 X_i$
$\hat{u}_i = Y_i - \hat{Y}_i = Y_i - \hat{b}_0 - \hat{b}_1 X_{1i} - \hat{b}_2 X_{2i} + \dots - \hat{b}_n X_{ni}$

Παραδοχές Κλασικού Γραμμικού Μοντέλου Παλινδρόμησης

Παραδοχή 1. Το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών.

Παραδοχή 2. Οι τιμές των ανεξάρτητων μεταβλητών είναι σταθερές σε επαναλαμβανόμενα δείγματα.

Παραδοχή 3. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, ο μέσος όρος της κατανομής των τιμών του διαταρακτικού όρου είναι μηδέν.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Παραδοχή 6. Αν τιμές των ανεξάρτητων μεταβλητών είναι στοχαστικές τότε είναι ανεξάρτητες (ή ασυσχέτιστες) από τις τιμές τους διαταρακτικού όρου.

Παραδοχή 7. Ο αριθμός των παρατηρήσεων του δείγματος πρέπει να είναι μεγαλύτερος από τον αριθμό των ανεξάρτητων μεταβλητών.

Παραδοχή 8. Οι τιμές των ανεξάρτητων μεταβλητών θα πρέπει να έχουν επαρκή μεταβλητότητα.

Παραδοχή 9. Το οικονομετρικό μοντέλο παλινδρόμησης θα πρέπει να είναι ορθά εξειδικευμένο.

Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Έλεγχος Εκτίμησης

- Η εκτίμηση του γραμμικού υποδείγματος πραγματοποιείται με τη χρήση ενός προγράμματος λογισμικού στατιστικής ανάλυσης (STATA).
- Τα αποτελέσματα της εκτίμησης του γραμμικού υποδείγματος θα πρέπει να ελεγχθούν ως προς τις ακόλουθες παραμέτρους:
 - Λογική συνέπεια ως προς τη θεωρία.
 - Στατιστική σημαντικότητα (για τις εκτιμηθείσες τιμές του σταθερού όρου και των συντελεστών παλινδρόμησης και για το σύνολο του υποδείγματος).
 - Ερμηνευτικότητα.
 - Παραβίαση παραδοχών.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Λογική Συνέπεια

- Το εκτιμηθέν υπόδειγμα θα πρέπει να ελεγχθεί για τη λογική συνέπειά του ως προς τη θεωρία και ειδικότερα ως προς τις ερευνητικές υποθέσεις που έχουν διατυπωθεί:
 - Οι εκτιμηθείσες τιμές των συντελεστών παλινδρόμησης που αφορούν τις σχετιζόμενες με τις ερευνητικές υποθέσεις ανεξάρτητες μεταβλητές έχουν το πρόσημο (θετικό ή αρνητικό) όπως προβλέπεται από τις αντίστοιχες ερευνητικές υποθέσεις;
 - Συνίσταται η κριτική ανάλυση με αντίστοιχα εμπειρικά αποτελέσματα προηγούμενων ερευνών.
 - Οι εκτιμηθείσες τιμές του σταθερού όρου και των συντελεστών παλινδρόμησης που αφορούν τις υπόλοιπες ανεξάρτητες μεταβλητές (ελέγχου) έχουν το πρόσημο (θετικό ή αρνητικό) όπως προβλέπεται από τις αντίστοιχες ερευνητικές υποθέσεις;
 - Συνίσταται η κριτική ανάλυση με αντίστοιχα εμπειρικά αποτελέσματα προηγούμενων ερευνών.

Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης: Λογική Συνέπεια

- Προσοχή στην ερμηνευτική. Σύμφωνα με την Παραδοχή 1 το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών. Ανάλογα με τη μαθηματική έκφραση των ανεξαρτήτων μεταβλητών διαφοροποιείται η ερμηνευτική για τις επιδράσεις των ανεξάρτητων μεταβλητών στη συμπεριφορά της εξαρτημένης μεταβλητής.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Λογική Συνέπεια

	Αρχική Εξειδίκευση	Γραμμικός Μετασχηματισμός	Ερμηνευτική Εκτιμηθέν Συντελεστή Παλίνδρομης b_1
Γραμμικό (Linear)	$Y_i = b_0 + b_1 X_i + u_i$	---	Η οριακή μεταβολή της Y όταν η X μεταβληθεί κατά μία μονάδα.
Λογαριθμικό – Γραμμικό (Log-Linear)	$Y_i = b_0 X_i^{b_1} e^{u_i}$	$\ln Y_i = b_0 + b_1 \ln X_i + u_i$	Η ποσοστιαία μεταβολή της Y για μία δεδομένη (σχετικά μικρή) ποσοστιαία μεταβολή της X (ελαστικότητα).
Ημιλογαριθμικό (Log – Lin)	$Y_i = Y_0 (1 + \lambda)^{X_i} u_i$	$\ln Y_i = \ln Y_0 + \ln(1 + \lambda) X_i + \ln u_i$ Θέτοντας $b_0 = \ln Y_0$ και $b_1 = \ln(1 + \lambda)$: $\ln Y_i = b_0 + b_1 X_i + \ln u_i$	Πολλαπλασιάζοντας τον εκτιμηθέν συντελεστή παλίνδρομης b_1 με 100 λαμβάνεται η ποσοστιαία μεταβολή της Y η X μεταβληθεί κατά μία μονάδα.
Ημιλογαριθμικό (Lin – Log)	$Y_i = b_0 + b_1 \ln X_i + \ln u_i$		Η μεταβολή της Y όταν η X μεταβληθεί κατά μία ποσοστιαία μονάδα.
Αντίστροφο (Reciprocal)	$Y_i = b_0 + b_1 \frac{1}{X_i} + u_i$	---	Ο εκτιμηθέν συντελεστή παλίνδρομης b_0 δίδει το όριο σύγκλισης και εκτιμηθέν συντελεστή παλίνδρομης b_1 το ρυθμό σύγκλισης

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Στατιστική Σημαντικότητα

- Το εκτιμηθέν υπόδειγμα θα πρέπει να ελεγχθεί για τη στατιστική σημαντικότητα των εκτιμηθεισών τιμών των συντελεστών παλινδρόμησης:
 - Οι εκτιμηθείσες τιμές των συντελεστών παλινδρόμησης που αφορούν τις σχετιζόμενες με τις ερευνητικές υποθέσεις ανεξάρτητες μεταβλητές (κατά κύριο λόγο) είναι στατιστικά σημαντικές;
 - Στατιστικά σημαντική είναι η εκτιμηθείσα τιμή του συντελεστή παλινδρόμησης αν αυτή είναι στατιστικά διάφορη του μηδενός. Αυτό γίνεται αντιληπτό αν η τιμή του p-value του στατιστικού t-student της εκτιμηθείσας τιμή του συντελεστή παλινδρόμησης είναι μικρότερη του 5% ή του 1% (ανάλογα).
 - Συνίσταται η κριτική ανάλυση με αντίστοιχα εμπειρικά αποτελέσματα προηγούμενων ερευνών.
 - Οι εκτιμηθείσες τιμές του σταθερού όρου και των συντελεστών παλινδρόμησης που αφορούν τις υπόλοιπες ανεξάρτητες μεταβλητές (ελέγχου) είναι στατιστικά σημαντικές;
 - Στατιστικά σημαντική είναι η εκτιμηθείσα τιμή του συντελεστή παλινδρόμησης αν αυτή είναι στατιστικά διάφορη του μηδενός. Αυτό γίνεται αντιληπτό αν η τιμή του p-value του στατιστικού t-student της εκτιμηθείσας τιμή του συντελεστή παλινδρόμησης είναι μικρότερη του 5% ή του 1% (ανάλογα).
 - Συνίσταται η κριτική ανάλυση με αντίστοιχα εμπειρικά αποτελέσματα προηγούμενων ερευνών.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Στατιστική Σημαντικότητα: Παράδειγμα

- Έστω ότι τα αποτελέσματα του υποδείγματος της παλινδρόμησης με εξαρτημένη μεταβλητή το μισθό (HWAGE) και ανεξάρτητες το επίπεδο εκπαίδευσης (EDUCATION), το φύλλο (GENDER), το προσδόκιμο ζωής (LFEXP), τη γεωγραφική περιοχή (REGION) και την ύπαρξη σωματείου εργαζομένων (UNION) έχει ως εξής :

Dependent Variable: HWAGE

Sample: 1 528

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.289796	1.258229	-3.409392	0.0007
EDUCATION	0.953006	0.082184	11.59596	0.0000
GENDER	-2.134171	0.391740	-5.447929	0.0000
LFEXP	0.104037	0.016888	6.160545	0.0000
REGION	-0.840832	0.427621	-1.966303	0.0498
UNION	1.427421	0.509978	2.798988	0.0053

Έλεγχος στατιστικής σημαντικότητας τιμών συντελεστών παλινδρόμησης.

R-squared	0.276707	Mean dependent var	9.047538
Adjusted R-squared	0.269779	S.D. dependent var	5.144082
S.E. of regression	4.395772	Akaike info criterion	5.810462
Sum squared resid	10086.51	Schwarz criterion	5.858974
Log likelihood	-1527.962	F-statistic	39.93978
Durbin-Watson stat	1.858629	Prob(F-statistic)	0.000000



Κλασσικό Γραμμικό Μοντέλο Παλινδρόμησης: Στατιστική Σημαντικότητα

- Το εκτιμηθέν υπόδειγμα θα πρέπει να ελεγχθεί για τη στατιστική σημαντικότητα του στο σύνολο του:
 - Μας ενδιαφέρει αν το εκτιμηθέν υπόδειγμα στο σύνολο του είναι στατιστικά σημαντικό.
 - Το σύνολο των εκτιμηθεισών τιμών όλων των συντελεστών παλινδρόμησης είναι από κοινού στατιστικά διάφορες του μηδενός. Αυτό γίνεται αντιληπτό αν η τιμή του p -value του στατιστικού F του υποδείγματος παλινδρόμησης είναι μικρότερη του 5% ή του 1% (ανάλογα) .

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Στατιστική Σημαντικότητα: Παράδειγμα

- Έστω ότι τα αποτελέσματα του υποδείγματος της παλινδρόμησης με εξαρτημένη μεταβλητή το μισθό (HWAGE) και ανεξάρτητες το επίπεδο εκπαίδευσης (EDUCATION), το φύλλο (GENDER), το προσδόκιμο ζωής (LFEXP), τη γεωγραφική περιοχή (REGION) και την ύπαρξη σωματείου εργαζομένων (UNION) έχει ως εξής :

Dependent Variable: HWAGE

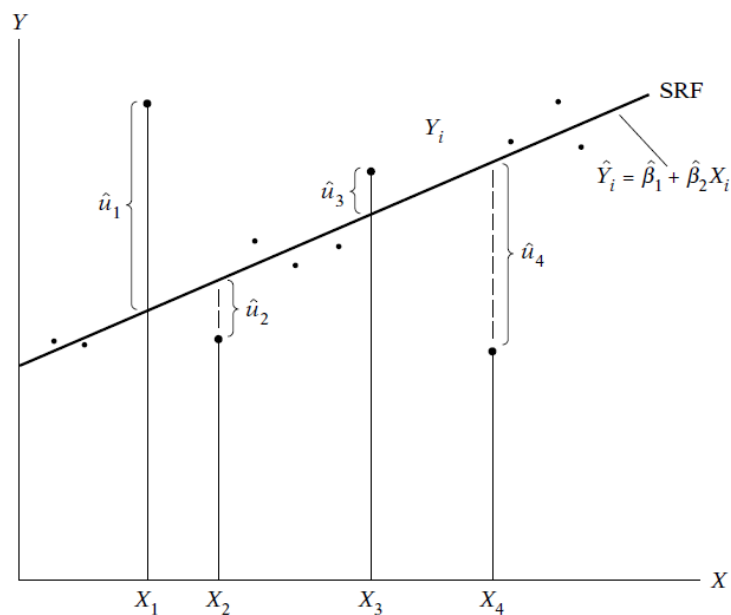
Sample: 1 528

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.289796	1.258229	-3.409392	0.0007
EDUCATION	0.953006	0.082184	11.59596	0.0000
GENDER	-2.134171	0.391740	-5.447929	0.0000
LFEXP	0.104037	0.016888	6.160545	0.0000
REGION	-0.840832	0.427621	-1.966303	0.0498
UNION	1.427421	0.509978	2.798988	0.0053
R-squared	0.276707	Mean dependent var	9.047538	
Adjusted R-squared	0.269779	S.D. dependent var	5.144082	
S.E. of regression	4.395772	Akaike info criterion	5.810462	
Sum squared resid	10086.51	Schwarz criterion	5.858974	
Log likelihood	-1527.962	F-statistic	39.93978	
Durbin-Watson stat	1.858629	Prob(F-statistic)	0.000000	

Έλεγχος στατιστικής σημαντικότητας υποδείγματος.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Ερμηνευτικότητα

- Το εκτιμηθέν υπόδειγμα θα πρέπει να ελεγχθεί αν παρέχει επαρκή ερμηνευτικότητα:
- **Ερμηνευτικότητα:** με δεδομένες τις εκτιμηθείσες τιμές των συντελεστών τι μέρος της συνολικής μεταβλητότητας της εξαρτημένης μεταβλητής ερμηνεύεται από τη μεταβλητότητα των ανεξάρτητων μεταβλητών;



- Στατιστικό R^2 (adjusted)

$$\bar{R}^2 = 1 - \frac{\sum \hat{u}_i^2 / (n - k)}{\sum y_i^2 / (n - 1)}$$

- Όσο τείνει προς την μονάδα (1) τόσο μεγαλύτερη είναι η ερμηνευτικότητα του εκτιμηθέντος υποδείγματος.

Κλασικό Γραμμικό Μοντέλο Παλινδρόμησης: Στατιστική Σημαντικότητα: Παράδειγμα

- Έστω ότι τα αποτελέσματα του υποδείγματος της παλινδρόμησης με εξαρτημένη μεταβλητή το μισθό (HWAGE) και ανεξάρτητες το επίπεδο εκπαίδευσης (EDUCATION), το φύλλο (GENDER), το προσδόκιμο ζωής (LFEXP), τη γεωγραφική περιοχή (REGION) και την ύπαρξη σωματείου εργαζομένων (UNION) έχει ως εξής :

Dependent Variable: HWAGE

Sample: 1 528

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-4.289796	1.258229	-3.409392	0.0007
EDUCATION	0.953006	0.082184	11.59596	0.0000
GENDER	-2.134171	0.391740	-5.447929	0.0000
LFEXP	0.104037	0.016888	6.160545	0.0000
REGION	-0.840832	0.427621	-1.966303	0.0498
UNION	1.427421	0.509978	2.798988	0.0053

R-squared	0.276707	Mean dependent var	9.047538
Adjusted R-squared	0.269779	S.D. dependent var	5.144082
S.E. of regression	4.395772	Akaike info criterion	5.810462
Sum squared resid	10086.51	Schwarz criterion	5.858974
Log likelihood	-1527.962	F-statistic	39.93978
Durbin-Watson stat	1.858629	Prob(F-statistic)	0.000000

Ερμηνευτικότητα υποδείγματος.

Υπόθεση Κανονικής Κατανομής

Παραδοχή 1. Το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών.

Παραδοχή 2. Οι τιμές των ανεξάρτητων μεταβλητών είναι σταθερές σε επαναλαμβανόμενα δείγματα.

Παραδοχή 3. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, ο μέσος όρος της κατανομής των τιμών του διαταρακτικού όρου είναι μηδέν.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Παραδοχή 6. Αν τιμές των ανεξάρτητων μεταβλητών είναι στοχαστικές τότε είναι ανεξάρτητες (ή ασυσχέτιστες) από τις τιμές τους διαταρακτικού όρου.

Παραδοχή 7. Ο αριθμός των παρατηρήσεων του δείγματος πρέπει να είναι μεγαλύτερος από τον αριθμό των ανεξάρτητων μεταβλητών.

Παραδοχή 8. Οι τιμές των ανεξάρτητων μεταβλητών θα πρέπει να έχουν επαρκή μεταβλητότητα.

Παραδοχή 9. Το οικονομετρικό μοντέλο παλινδρόμησης θα πρέπει να είναι ορθά εξειδικευμένο.

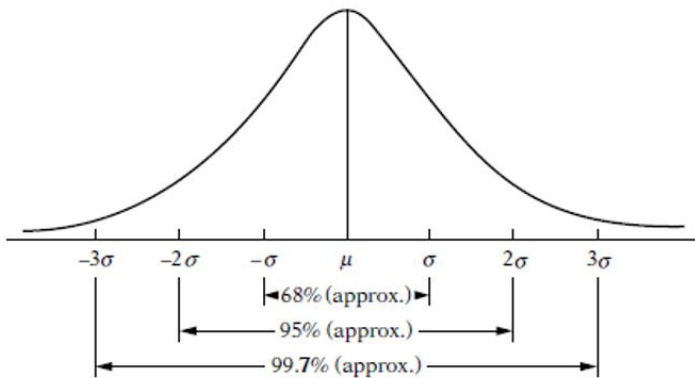
Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Υπόθεση Κανονικής Κατανομής

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad -\infty < x < \infty$$

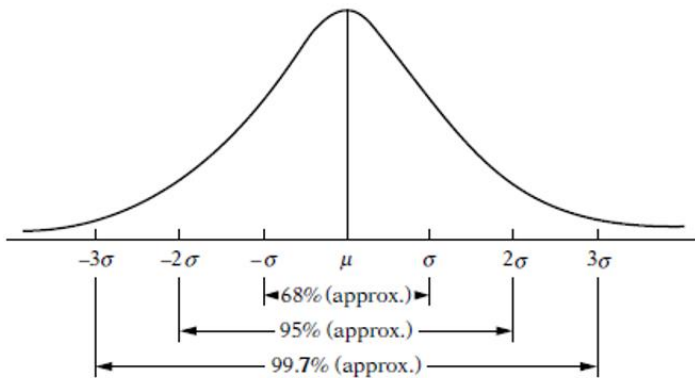


Areas under the normal curve.

- **Παραδοχή 11:**
 - Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.
 - Στο διπλανό διάγραμμα εμφανίζεται το γράφημα της τυποποιημένης κανικής κατανομής.
- **Αιτίες:**
 - Σφάλμα εξειδίκευσης του υποδείγματος.
 - Εσφαλμένος ορισμός μεταβλητών.
 - Εσφαλμένη διαδικασία συλλογής δεδομένων.
 - Μικρό δείγμα.
- **Συνέπειες αυτοσυσχέτισης:**
 - Οι εκτιμημένες τιμές των συντελεστών παλινδρόμησης είναι αμερόληπτες και αποτελεσματικές ΑΛΛΑ αν οι τιμές του διαταρακτικού όρου δεν ακολουθούν την κανονική κατανομή δημιουργούνται θέματα στον ορισμό των διαστημάτων εμπιστοσύνης.

Υπόθεση Κανονικής Κατανομής

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad -\infty < x < \infty$$

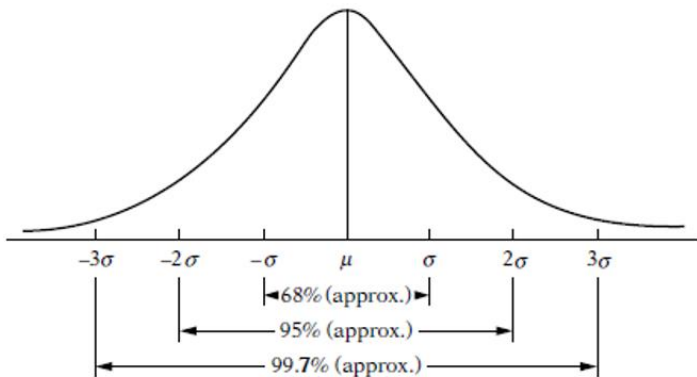


Areas under the normal curve.

- Διαγνωστικοί ελέγχοι (σύμφωνα με τη βιβλιογραφία):
 - Ιστόγραμμα των τιμών του διαταρακτικού όρου.
 - Anderson–Darling normality test (A2 statistic).
 - Jarque–Bera (JB) test of normality.
- Αντιμετώπιση (σύμφωνα με τη βιβλιογραφία):
 - Καταλληλότερη η εκτίμηση του υποδείγματος με μέγιστη πιθανοφάνειας αντί της μεθόδου ελαχιστοποίησης των τετραγώνων των αποκλίσεων.
 - Αύξηση του μεγέθους τους δείγματος.

Έλεγχος Κανονικής Κατανομής με το Stata

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad -\infty < x < \infty$$



Areas under the normal curve.

Διάγνωση προβλήματος - Εντολές Stata:

Εντολές swilk, sfrancia και sktest.

Σύνταξη εντολής (swilk):

```
xtreg vrY vrX1 vrX2 . . . vrXn
predict e, residuals
swilk e
```

Σύνταξη εντολής (sfrancia):

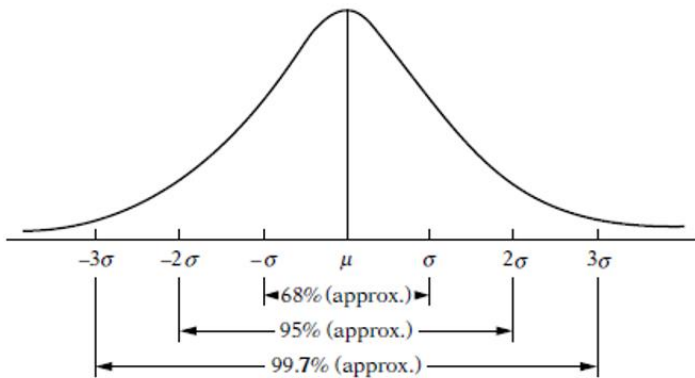
```
xtreg vrY vrX1 vrX2 . . . vrXn
predict e, residuals
sfrancia e
```

Σύνταξη εντολής (sktest):

```
xtreg vrY vrX1 vrX2 . . . vrXn
predict e, residuals
sktest e
```

Έλεγχος Κανονικής Κατανομής με το Stata

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right) \quad -\infty < x < \infty$$



Areas under the normal curve.

- Ερμηνεία αποτελεσμάτων:

- H_0 : Κανονική
- `swilk`, `sfrancia`: H_0 απορρίπτεται αν το p-value του στατιστικού z είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>z= XXX < 0.05`).
- `sktest`: η H_0 απορρίπτεται αν το p-value του στατιστικού z είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>z= XXX < 0.05`).

Πολυσυγγραμμικότητα

Παραδοχή 1. Το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών.

Παραδοχή 2. Οι τιμές των ανεξάρτητων μεταβλητών είναι σταθερές σε επαναλαμβανόμενα δείγματα.

Παραδοχή 3. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, ο μέσος όρος της κατανομής των τιμών του διαταρακτικού όρου είναι μηδέν.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Παραδοχή 6. Αν τιμές των ανεξάρτητων μεταβλητών είναι στοχαστικές τότε είναι ανεξάρτητες (ή ασυσχέτιστες) από τις τιμές τους διαταρακτικού όρου.

Παραδοχή 7. Ο αριθμός των παρατηρήσεων του δείγματος πρέπει να είναι μεγαλύτερος από τον αριθμό των ανεξάρτητων μεταβλητών.

Παραδοχή 8. Οι τιμές των ανεξάρτητων μεταβλητών θα πρέπει να έχουν επαρκή μεταβλητότητα.

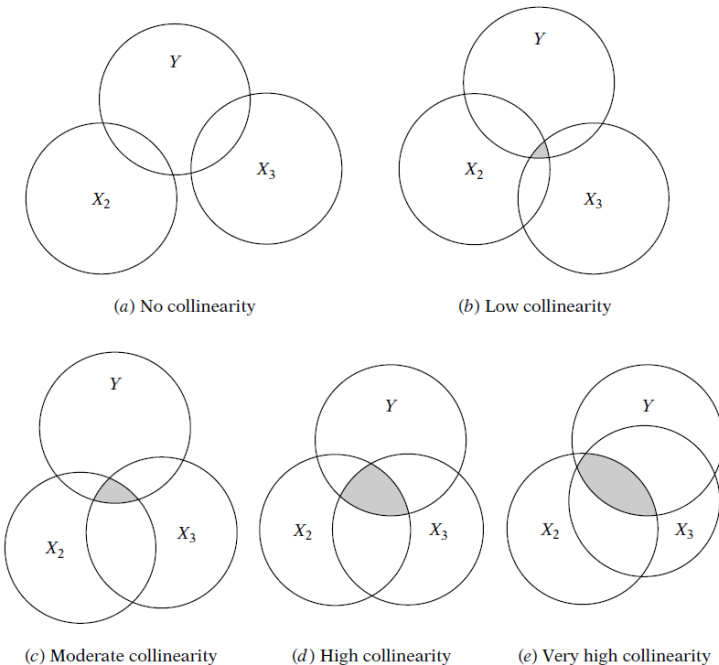
Παραδοχή 9. Το οικονομετρικό μοντέλο παλινδρόμησης θα πρέπει να είναι ορθά εξειδικευμένο.

Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

Πολυσυγγραμμικότητα



- Παραδοχή 10:
 - Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση
 - Στο διπλανό διάγραμμα μόνο η περίπτωση (a) ικανοποιεί την παραδοχή 10 πλήρως
- Αιτίες:
 - Σφάλμα εξειδίκευσης του υποδείγματος και ειδικότερα εσφαλμένη μαθηματική εξειδίκευση του οικονομετρικού μοντέλου ή πλεονασμός ανεξάρτητων μεταβλητών στο υπόδειγμα.
 - Εσφαλμένος ορισμός μεταβλητών.
 - Εσφαλμένη διαδικασία συλλογής δεδομένων.
- Συνέπειες αυτοσυσχέτισης:
 - Οι εκτιμημένες τιμές των συντελεστών παλινδρόμησης είναι στατιστικά ασήμαντες ενός συντελεστή προσδιορισμού R^2 είναι μεγάλος.
 - Έλλειψη ακρίβειας στην εκτίμηση των τιμών των συντελεστών παλινδρόμησης.
 - Οι εκτιμημένες τιμές των συντελεστών παλινδρόμησης είναι αμερόληπτες αλλά δεν είναι αποτελεσματικές, δηλαδή η τυπική απόκλισή τους δεν είναι η μικρότερη δυνατή.

Έλεγχος και Αντιμετώπιση Πολυσυγγραμικότητας με το Stata

- Διάγνωση προβλήματος - Εντολές Stata:

- Εντολή vif.
- Σύνταξη εντολής (estat bgodfrey):

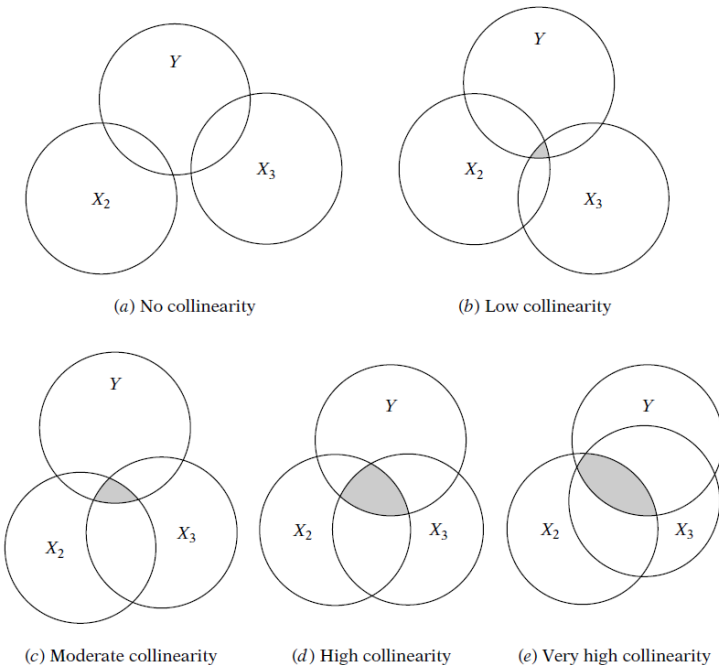
```
reg vrY vrX1 vrX2 . . . vrXn
vif
```

- Ερμηνεία αποτελεσμάτων:

- Ένα $VIF > 10$ ή ένα $1/VIF < 0,10$ υποδηλώνει την παρουσία πολυσυγγραμικότητας στο μοντέλο

- Αντιμετώπιση προβλήματος – Εντολές Stata:

- Συνήθως το Stata αφαιρεί αυτόματα τη μεταβλητή που προκαλεί πολυσυγγραμικότητα.
- Αξιολογική κρίση: αποτελεό δομικό πρόβλημα ορισμού του μοντέλου και θα πρέπει να αντιμετωπισθεί σε επίπεδο μαθηματικής εξειδίκευσης του μοντέλου γραμμικής παλινδρόμησης.



Ετεροσκεδαστικότητα

Παραδοχή 1. Το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών.

Παραδοχή 2. Οι τιμές των ανεξάρτητων μεταβλητών είναι σταθερές σε επαναλαμβανόμενα δείγματα.

Παραδοχή 3. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, ο μέσος όρος της κατανομής των τιμών του διαταρακτικού όρου είναι μηδέν.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Παραδοχή 6. Αν τιμές των ανεξάρτητων μεταβλητών είναι στοχαστικές τότε είναι ανεξάρτητες (ή ασυσχέτιστες) από τις τιμές τους διαταρακτικού όρου.

Παραδοχή 7. Ο αριθμός των παρατηρήσεων του δείγματος πρέπει να είναι μεγαλύτερος από τον αριθμό των ανεξάρτητων μεταβλητών.

Παραδοχή 8. Οι τιμές των ανεξάρτητων μεταβλητών θα πρέπει να έχουν επαρκή μεταβλητότητα.

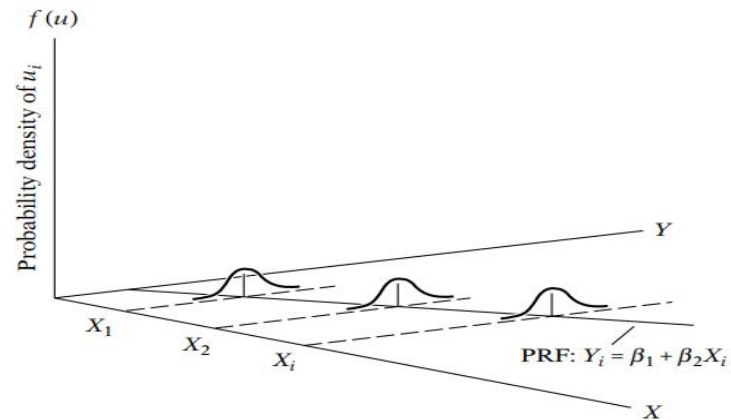
Παραδοχή 9. Το οικονομετρικό μοντέλο παλινδρόμησης θα πρέπει να είναι ορθά εξειδικευμένο.

Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

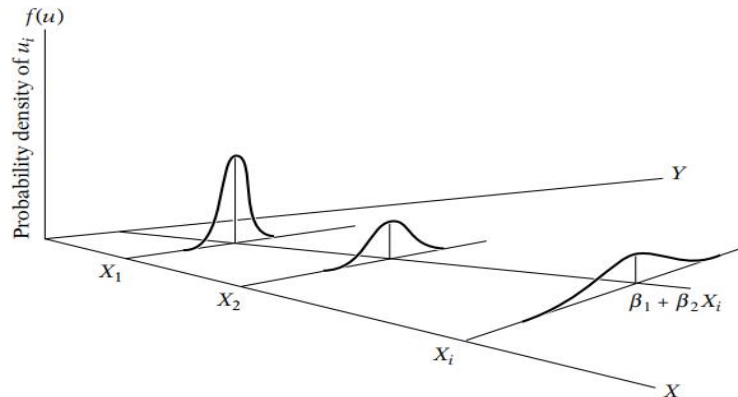
Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Ετεροσκεδαστικότητα



Homoscedasticity.



Heteroscedasticity.

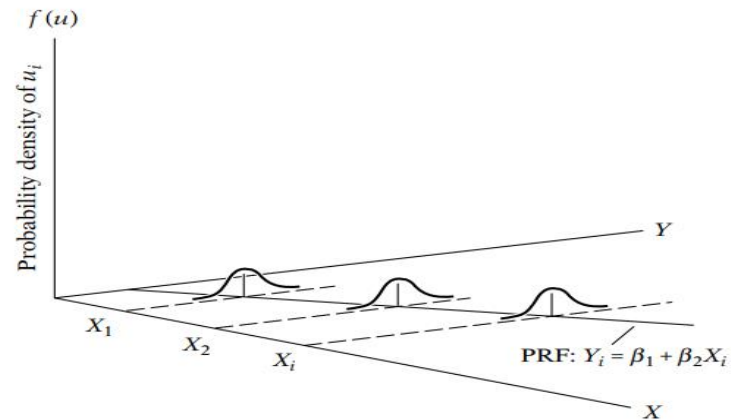
Παραδοχή 4:

- Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).
- Αν η παραπάνω παραδοχή δεν ισχύει τότε εμφανίζεται το πρόβλημα της ετεροσκεδαστικότητας.

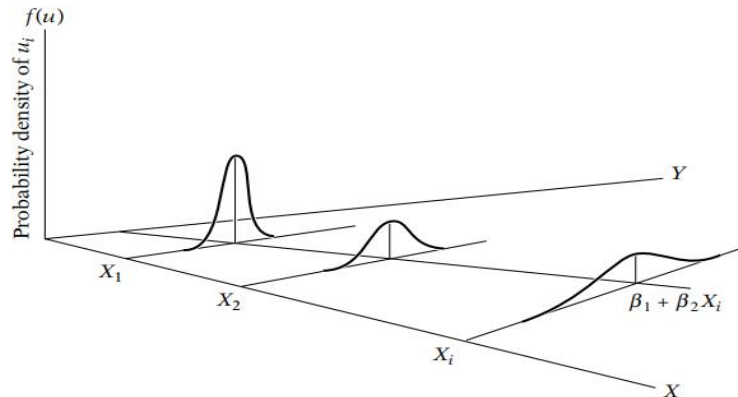
Αιτίες:

- Βελτίωση της διαδικασίας συλλογής δεδομένων (μείωση σφαλμάτων και μείωση διακύμανσης).
- Ύπαρξη ακραίων τιμών στο δείγμα παρατηρήσεων.
- Εσφαλμένη εξειδίκευση του υποδείγματος.
- Ιδιαιτερότητες στη συμπεριφορά μίας ή περισσότερων ανεξαρτήτων μεταβλητών (π.χ., έντονη ασυμμετρία).
- Ανεπάρκεια δεδομένων ή εσφαλμένος μετασχηματισμός τους.
- Παρατηρείται συχνότερα σε διαστρωματικά δεδομένα συγκριτικά με παρατηρήσεις μία χρονολογική σειράς.

Ετεροσκεδαστικότητα



Homoscedasticity.



Heteroscedasticity.

Συνέπειες ετεροσκεδαστικότητας:

- Οι εκτιμημένες τιμές των συντελεστών παλινδρόμησης είναι αμερόληπτες αλλά δεν είναι αποτελεσματικές, δηλαδή η τυπική απόκλισή τους δεν είναι η μικρότερη δυνατή.

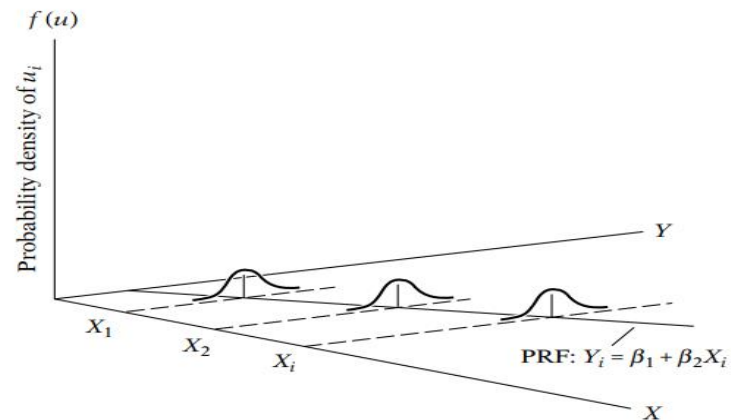
Διαγνωστικοί ελέγχοι (σύμφωνα με τη βιβλιογραφία):

- Park test, Glejser test, Spearman's rank correlation test.
- Goldfeld-Quandt test.
- Breusch-Pagan-Godfrey test.
- White's General heteroscedasticity test.

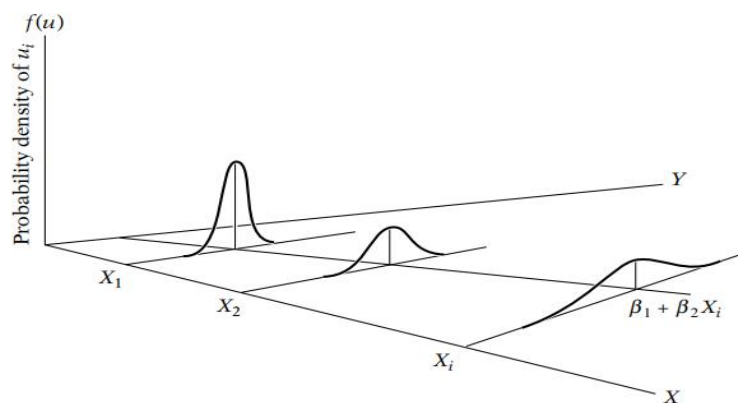
Αντιμετώπιση (σύμφωνα με τη βιβλιογραφία):

- Αν είναι γνωστή η πραγματική διακύμανση (σ^2) της κατανομής του διαταρακτικού όρου τότε εκτελείται εκτίμηση του οικονομετρικού μοντέλου με την τεχνική Weighted Least Squares.
- Αν δεν είναι γνωστή η πραγματική διακύμανση (σ^2) της κατανομής του διαταρακτικού όρου τότε γίνεται χρήση διακυμάνσεων και τυπικών σφαλμάτων προσαρμοσμένων κατά White (White's Heteroscedasticity-Consistent Variances and Standard Errors).

Έλεγχος και Αντιμετώπιση Ετεροσκεδαστικότητας με το Stata



Homoscedasticity.



Heteroscedasticity.

Διάγνωση προβλήματος - Εντολές Stata:

- Εντολές `estat hettest` και `estat imtest, white`.
- Σύνταξη εντολής (`estat hettest`):
`reg vrY vrX1 vrX2 . . . vrXn`
`estat hettest`
- Σύνταξη εντολής (`estat imtest, white`):
`reg vrY vrX1 vrX2 . . . vrXn`
`estat imtest, white`

Ερμηνεία αποτελεσμάτων:

- H_0 : Ομοσκεδαστικότητα
- Η H_0 απορρίπτεται αν το p-value του στατιστικού χ^2 είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>chi2 = XXX < 0.05`).

Αντιμετώπιση προβλήματος – Εντολές Stata:

- Επιλογή `robust`
- Σύνταξη:
`reg vrY vrX1 vrX2 . . . vrXn, robust`

Αυτοσυσχέτιση

Παραδοχή 1. Το μοντέλο παλινδρόμησης είναι γραμμικό ως προς τις εκτιμημένες τιμές των συντελεστών παλινδρόμησης άσχετα από τη μαθηματική έκφραση των ανεξάρτητων μεταβλητών.

Παραδοχή 2. Οι τιμές των ανεξάρτητων μεταβλητών είναι σταθερές σε επαναλαμβανόμενα δείγματα.

Παραδοχή 3. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, ο μέσος όρος της κατανομής των τιμών του διαταρακτικού όρου είναι μηδέν.

Παραδοχή 4. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, η διακύμανση της κατανομής των τιμών του διαταρακτικού όρου είναι σταθερή (ομοσκεδαστικότητα).

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Παραδοχή 6. Αν τιμές των ανεξάρτητων μεταβλητών είναι στοχαστικές τότε είναι ανεξάρτητες (ή ασυσχέτιστες) από τις τιμές τους διαταρακτικού όρου.

Παραδοχή 7. Ο αριθμός των παρατηρήσεων του δείγματος πρέπει να είναι μεγαλύτερος από τον αριθμό των ανεξάρτητων μεταβλητών.

Παραδοχή 8. Οι τιμές των ανεξάρτητων μεταβλητών θα πρέπει να έχουν επαρκή μεταβλητότητα.

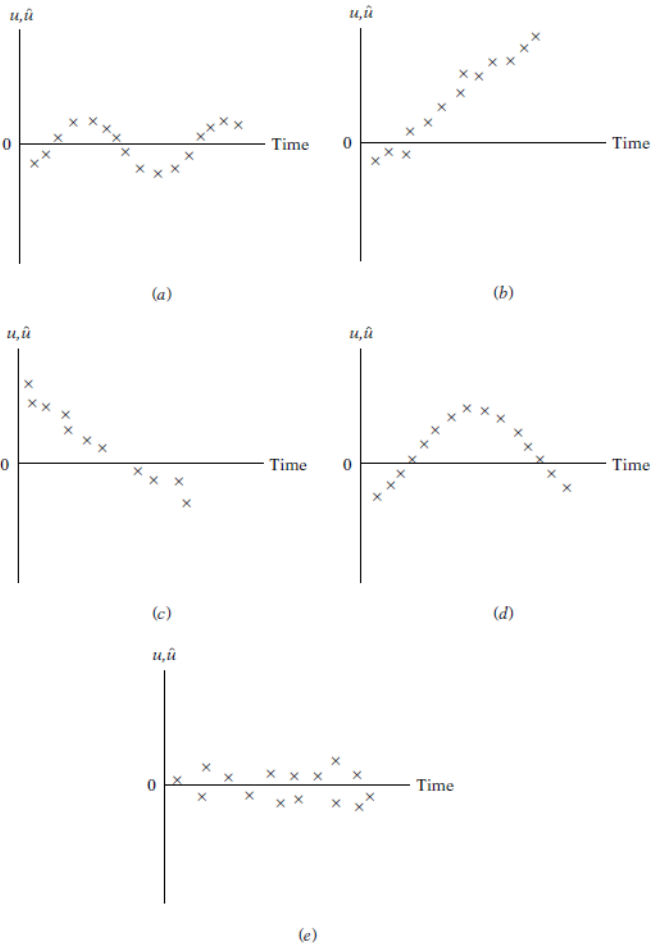
Παραδοχή 9. Το οικονομετρικό μοντέλο παλινδρόμησης θα πρέπει να είναι ορθά εξειδικευμένο.

Παραδοχή 10. Μεταξύ των ανεξάρτητων μεταβλητών δεν πρέπει να υπάρχει γραμμική συσχέτιση.

Παραδοχή 11. Οι τιμές του διαταρακτικού όρου πρέπει να ακολουθούν την κανονική κατανομή.

Παραδοχή 5. Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).

Αυτοσυσχέτιση



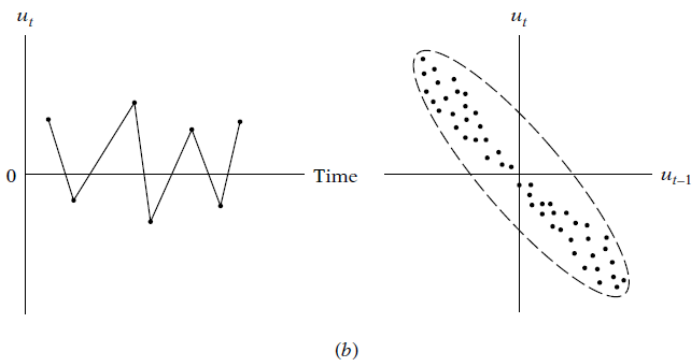
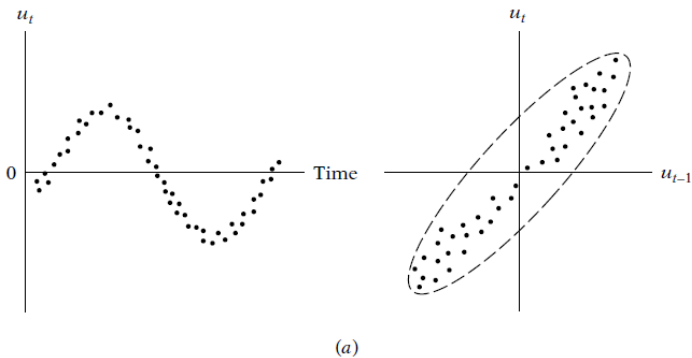
• Παραδοχή 5:

- Για δεδομένες τιμές των ανεξάρτητων μεταβλητών, οι τιμές του διαταρακτικού όρου είναι μηδέν δεν συσχετίζονται μεταξύ τους (απουσία αυτοσυσχέτισης).
- Στο διπλανό διάγραμμα μόνο η περίπτωση (e) ικανοποιεί την παραδοχή 5 και άρα δεν υπάρχει αυτοσυσχέτιση.

• Αιτίες:

- Αδράνεια: γενικευμένη διαχρονική τάση οικονομικών μεταβλητών προς μια κατεύθυνση.
- Σφάλμα εξειδίκευσης του υποδείγματος και ειδικότερα παράληψη ανεξάρτητης μεταβλητής (ως εκ τούτου ο διαταρακτικός όρος συμπεριλαμβάνει τις μεταβολές της εσφαλμένα παραληφθείσας ανεξάρτητης μεταβλητής).
- Σφάλμα εξειδίκευσης του υποδείγματος και ειδικότερα εσφαλμένη μαθηματική εξειδίκευση του οικονομετρικού μοντέλου (ως εκ τούτου ο διαταρακτικός όρος συμπεριλαμβάνει τις μεταβολές ανεξάρτητων μεταβλητών κατά το μέρος της εξειδίκευσης που παραλήφθηκε).
- Αυτοσυσχέτιση της εξαρτημένης μεταβλητής με παρελθούσες τιμές της η οποία έχει αγνοηθεί κατά την εξειδίκευση του οικονομετρικού μοντέλου.

Αυτοσυσχέτιση

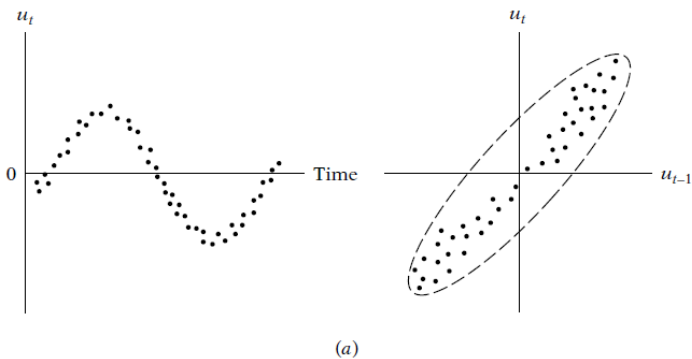


(a) Positive and (b) negative autocorrelation.

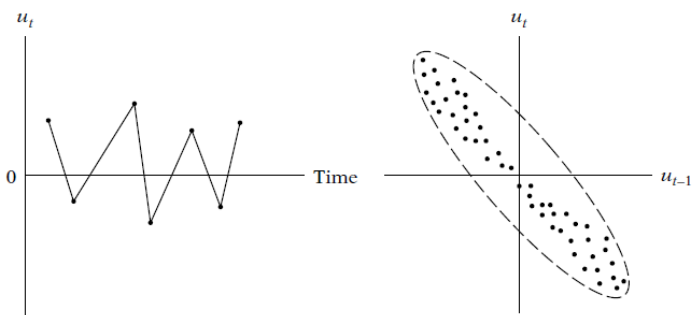
- Αιτίες (συνέχεια):

- Αυτοσυσχέτιση της εξαρτημένης μεταβλητής με παρελθούσες τιμές της η οποία έχει αγνοηθεί κατά την εξειδίκευση του οικονομετρικού μοντέλου.
- Μόχλευση δεδομένων.
- Μετασχηματισμός δεδομένων μέσω υπολογιστικών διαδικασιών για τον υπολογισμό των μεταβλητών.
- Μη στάσιμες μεταβλητές (χρονολογικές σειρές).

Αυτοσυσχέτιση



(a)

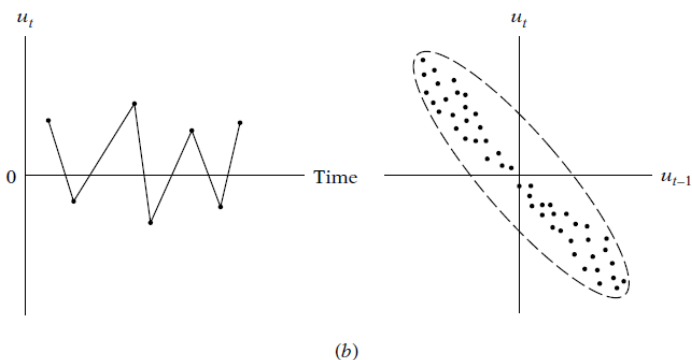
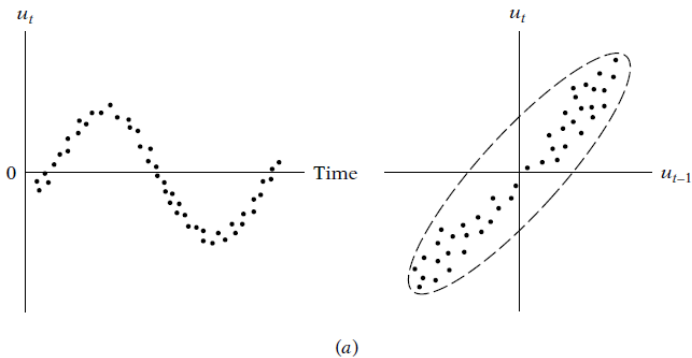


(b)

(a) Positive and (b) negative autocorrelation.

- **Συνέπειες αυτοσυσχέτισης:**
 - Οι εκτιμημένες τιμές των συντελεστών παλινδρόμησης είναι αμερόληπτες αλλά δεν είναι αποτελεσματικές, δηλαδή η τυπική απόκλισή τους δεν είναι η μικρότερη δυνατή.
- **Διαγνωστικοί ελέγχοι (σύμφωνα με τη βιβλιογραφία):**
 - Διαχρονικό γράφημα των τιμών του διαταρακτικού όρου.
 - Durbin-Watson d test.
 - Durbin-Watson h test
 - Breusch–Godfrey test.
- **Αντιμετώπιση (σύμφωνα με τη βιβλιογραφία):**
 - Εξετάζεται το ενδεχόμενο το οικονομετρικό μοντέλο να μην έχει την κατάλληλη εξειδίκευση και λαμβάνονται κατάλληλες ενέργειες για τη διόρθωσή του.
 - Εκτίμηση του οικονομετρικού υποδείγματος με τη χρήση μεθόδου Generalized Least Squares – GLS (=θεωρείται ότι ο διαταρακτικός όρος ε_t ορίζεται ως $\varepsilon_t = \rho \varepsilon_{t-1} + u_t$ με $-1 < \rho < 1$) ή Feasible Generalized Least Squares – FGLS.
 - Χρήση διακυμάνσεων και τυπικών σφαλμάτων προσαρμοσμένων για ετεροσκεδαστικότητα και αυτοσυσχέτιση (HAC - (heteroscedasticity- and autocorrelation-consistent) - Newey–West.

Έλεγχος και Αντιμετώπιση Αυτοσυσχέτισης με το Stata



(a) Positive and (b) negative autocorrelation.

- Διάγνωση προβλήματος - Εντολές Stata:

- Εντολή `estat bgodfrey`.
- Σύνταξη εντολής (`estat bgodfrey`):

```
generate time = _n
tsset time
reg vrY vrX1 vrX2 . . . vrXn
estat bgodfrey
```

- Ερμηνεία αποτελεσμάτων:

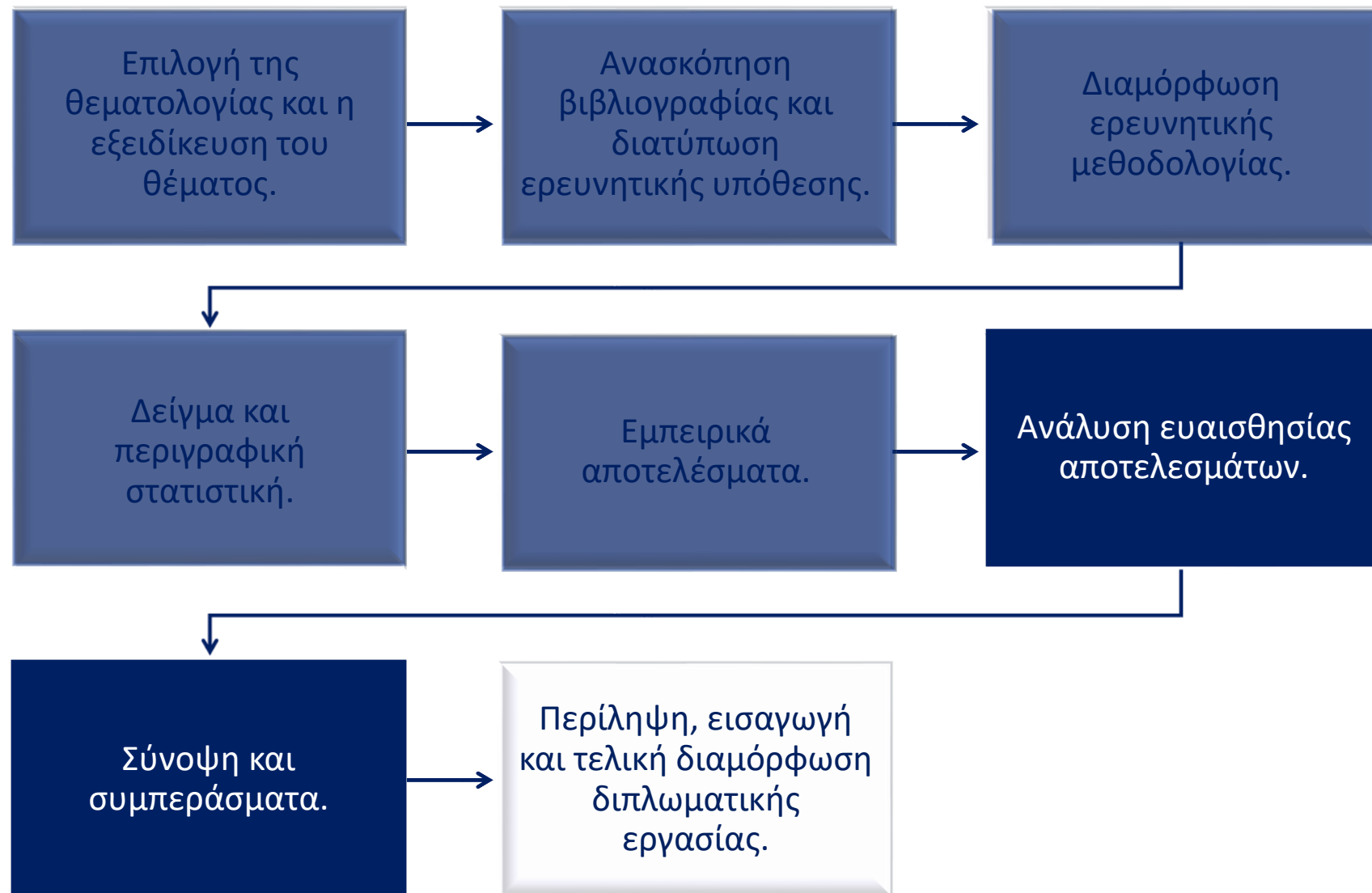
- H_0 : Δεν υπάρχει αυτοσυσχέτιση
- Η H_0 απορρίπτεται αν το p-value του στατιστικού χ^2 είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>chi2 = XXX < 0.05`).

- Αντιμετώπιση προβλήματος – Εντολές Stata:

- Δημιουργία αριθμητικής μεταβλητής `Firm_id` η οποία αναφέρεται στο μοναδικό κωδικό που διακρίνεται η κάθε εταιρεία στο δείγμα των δεδομένων (αν υπάρχει ήδη μία τέτοια μεταβλητή τότε χρησιμοποιούμε αυτή τη μεταβλητή στη θέση της `Firm_id`)
- Επιλογή `cluster(Firm_id)`
- Σύνταξη:

```
reg vrY vrX1 vrX2 . . . vrXn, cluster(Firm_id)
```

Ανάλυση Ευαισθησίας Αποτελεσμάτων Συμπεράσματα και Σύνοψη



Ανάλυση Ευαισθησίας Αποτελεσμάτων

- Στο στάδιο αυτό της διπλωματικής εργασίας λαμβάνει χώρα η ανάλυση ευαισθησίας των εμπειρικών αποτελεσμάτων.
- Η ανάγκη για ανάλυση ευαισθησίας αναδύεται ακόμη και αν τα κύρια εμπειρικά αποτελέσματα επιβεβαιώνουν τις ερευνητικές υποθέσεις περαιτέρω έλεγχοι για μία σειρά από λόγους:
 - Τα εμπειρικά αποτελέσματα επιδέχονται εναλλακτικής ερμηνείας.
 - Ανάγκη για αποκλεισμό εναλλακτικής ερμηνείας (δηλαδή ερμηνείας διαφορετικής από αυτή που προβλέπουν οι ερευνητικές υποθέσεις) των εμπειρικών αποτελεσμάτων.
 - Ανάγκη για επιβεβαίωση κύριων ερευνητικών αποτελεσμάτων με εναλλακτική μεθοδολογία, μεταβλητές, εμπειρικό δείγμα, τεχνικές εκτίμησης του γραμμικού μοντέλου παλινδρόμησης, κ.λπ.
- Αποτελεί εξειδικευμένο θέμα οικονομετρίας. Για λόγους γενικής αντίληψης θα παρουσιασθούν οι έλεγχος Hausman και ο έλεγχος Ramsey RESET.

Έλεγχος Hausman

- **Δεδομένα τύπου πάνελ:** συνδυασμός χρονολογικής σειράς και διαστρωματικών δεδομένων, δηλαδή ένας αριθμός χρονολογικών ακολουθιών τιμών και έναν αριθμό περιπτώσεων. Παραδείγματα: το ύψος των πωλήσεων ενός αριθμού εταιρειών της τελευταίας δεκαετίας, η αξία των μηναίων πωλήσεων ενός αριθμού εταιρειών για το χρονικό διάστημα από το 1990 έως και το 2020, το βάρος ενός αριθμού αθλητών τους τελευταίους 36 μήνες.
 - Η κάθε περίπτωση (εταιρεία) έχει ιδιαίτερα χαρακτηριστικά τα οποία αν δεν ληφθούν υπόψη κατά τη διαμόρφωση του υποδείγματος παλινδρόμησης προκύπτει σφάλμα εξειδίκευσης.
 - Τα χαρακτηριστικά αυτά αντιπροσωπεύονται δυνητικά από ένα σύνολο μεταβλητών.
 - Η παράλειψη τους σημαίνει ότι εμπεριέχονται στον διαταρακτικό όρο και εν συνεχεία αυτές (ΚΑΙ ΑΡΑ ΚΑΙ ο διαταρακτικός όρος) να συσχετίζονται με τις ανεξάρτητες μεταβλητές (παραβιάζεται η παραδοχή 6).

Έλεγχος Hausman

- Λύση Α: Να εμπλουτισθεί το υπόδειγμα με ψευδομεταβλητές (δίτιμες) που υποδηλώνουν τις επιμέρους περιπτώσεις.
 - Πιθανό πρόβλημα πολυσυγγραμικότητα και επιπλέον πρακτικά δύσκολο όταν ο αριθμός των περιπτώσεων είναι πολύ μεγάλος.
- Λύση Β: Κατά την εκτίμηση του υποδείγματος παλινδρόμησης να θεωρηθεί ότι ο διαταρακτικός όρος ότι είναι συνάρτηση σταθερών (fixed effects) ή τυχαίων (random effects) επιδράσεων των (σταθερών στο χρόνο) χαρακτηριστικών των περιπτώσεων.
 - Κρίσιμο ερώτημα: Ποια υπόθεση θα πρέπει να υιοθετηθεί; Σταθερές (fixed effects) ή τυχαίες (random effects) επιδράσεις;
 - Απάντηση έλεγχος Hausman

Έλεγχος Hausman: Stata

- Έστω ότι επιθυμούμε να εκτιμήσουμε ένα υπόδειγμα (πολυμεταβλητής) γραμμικής παλινδρόμησης με δεδομένα πάνελ που αφορά μια εξαρτημένη μεταβλητή (vrY) και ένα πλήθος ανεξάρτητων μεταβλητών (έστω τρεις με την ονομασία $vrX1$, $vrX2$ και $vrX3$ αντίστοιχα) με δεδομένα πάνελ, και με τυχαίες επιδράσεις.
- Για να εξάγουμε στατιστικό συμπέρασμα αν θα επιλέξουμε στην εκτίμηση μας σταθερές επιδράσεις (επιλογή fe) ή τυχαίες επιδράσεις (επιλογή re) εκτελούμε τον έλεγχο `hausman`.
- Η ακολουθία των εντολών στο STATA είναι η εξής:

```
xtreg vrY vrX1 vrX2 vrX3, fe  
estimates store fixed  
xtreg vrY vrX1 vrX2 vrX3, re  
estimates store random  
hausman fixed random
```
- Η μηδενική υπόθεση είναι ότι πρέπει να επιλεχθούν οι τυχαίες επιδράσεις απορρίπτεται (δεν γίνεται αποδεκτή) αν το p -value του στατιστικού χ^2 είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>chi2 = XXX < 0.05`).

Έλεγχος Ramsey RESET

- Μία συνηθισμένη ερευνητική ανησυχία είναι η εσφαλμένη εξειδίκευση της μαθηματικής έκφρασης της εξίσωσης παλινδρόμησης. Μία από τις πιθανές αιτίες είναι η παράλειψη ανεξάρτητης μεταβλητής (omitted-variable bias).
- Ένας συνήθης έλεγχος για τη διάγνωση του προβλήματος εσφαλμένης εξειδίκευσης εξαιτίας της παράλειψης ανεξάρτητης μεταβλητής είναι ο έλεγχος Ramsey RESET.
- Η εντολή για τον έλεγχο Ramsey RESET είναι η εντολή `ovtest` και εκτελείται μετά την εκτέλεση της εντολής `regress` (στην επόμενη γραμμή του πεδίου των εντολών του STATA). Η ακολουθία εντολών στο STATA θα έχει ως εξής:

```
regress vrY vrX1  
ovtest
```
- Η μηδενική υπόθεση είναι ότι δεν έχουμε σφάλμα παράλειψης μεταβλητής απορρίπτεται (δεν γίνεται αποδεκτή) αν το p -value του στατιστικού X^2 είναι μικρότερο του 5% (δηλαδή θα δείτε στο report το εξής: `prob>F = XXX < 0.05`).

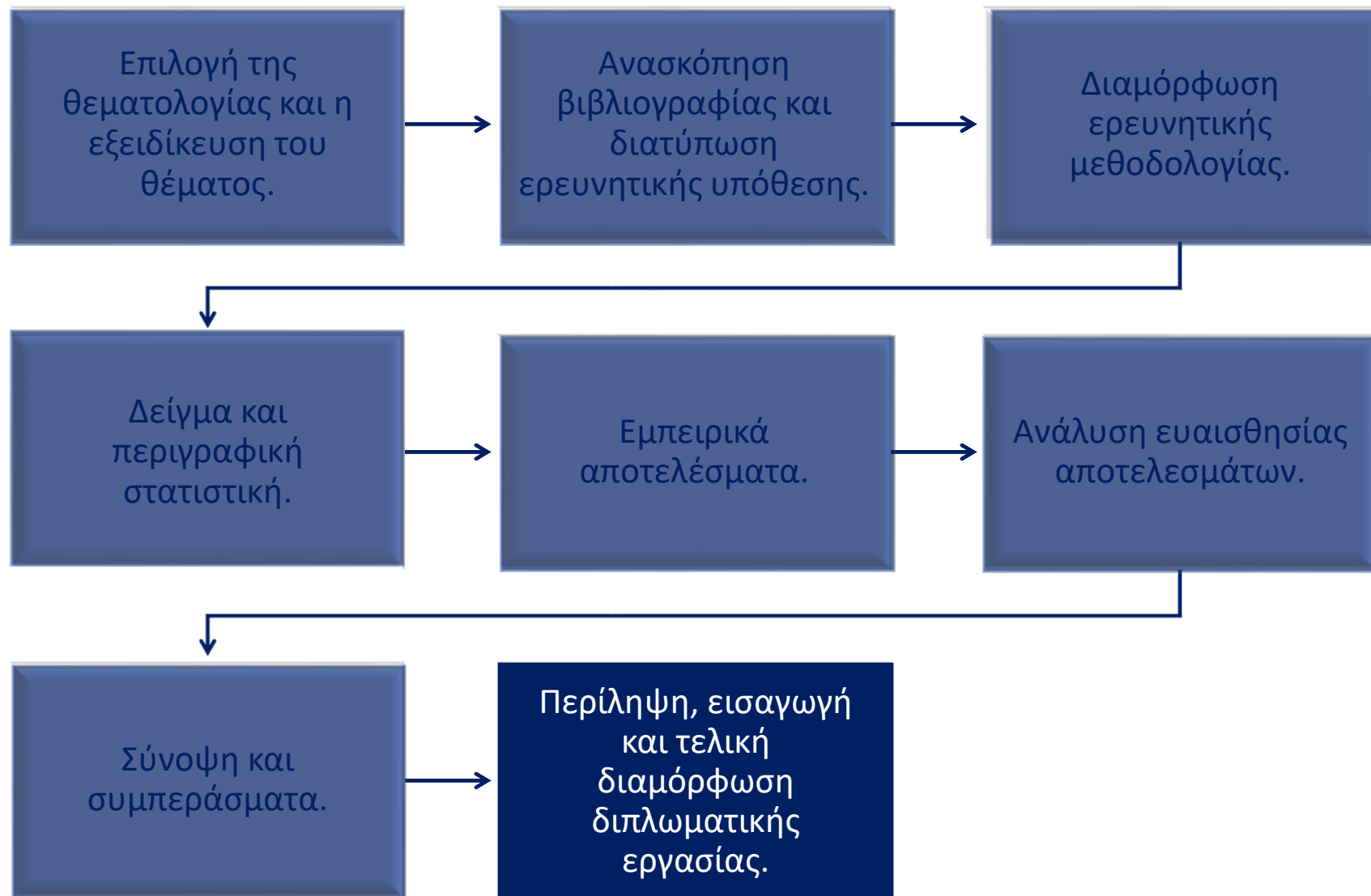
Σύνοψη και Συμπεράσματα

- Σε αυτή την ενότητα, πραγματοποιείται μια σύνοψη της διπλωματικής εργασίας και των βασικών συμπερασμάτων που προέκυψαν.
- Πρόκειται για ιδιαίτερα σημαντική ενότητα καθώς είναι αυτή που θα μείνει πιο έντονα στον αναγνώστη. Θα πρέπει να παρουσιαστεί με σύντομο τρόπο βασικότερα σημεία της διπλωματικής εργασίας, δηλαδή ο στόχος, το κίνητρο, το δείγμα, τα βασικά αποτελέσματα και τα συμπεράσματα που προέκυψαν.
- Σημαντικό σημείο είναι η αναγνώριση τυχόν περιορισμών της διπλωματικής εργασίας, όπως περιορισμοί σε διαθέσιμα στοιχεία, περιορισμοί της μεθοδολογίας που χρησιμοποιήθηκε, κ.λπ.
- Τέλος, αναγνωρίζονται τυχόν ευκαιρίες για περαιτέρω διερεύνηση και μελλοντική έρευνα.

Ανάλυση Γραμμικής Παλινδρόμησης

Εφαρμογή 6^η

Περίληψη, Εισαγωγή και Τελική Διαμόρφωση Διπλωματικής Εργασίας





Περίληψη

- Η περίληψη αποτελεί μια πολύ σύντομη παρουσίαση της διπλωματικής εργασίας.
- Σκοπός είναι με έναν ελάχιστο αριθμό λέξεων (έως 300 λέξεις) να διατυπωθεί ξεκάθαρα ο σκοπός της εργασίας, το πλαίσιο, τα βασικά ευρήματα και τα σημαντικότερα συμπεράσματα.



Εισαγωγή

- Η εισαγωγή αποτελεί το πρώτο μέρος της διπλωματικής εργασίας.
- Στην εισαγωγή, ο φοιτητής θα πρέπει να αναφέρει το θέμα προς διερεύνηση και το ερευνητικό κίνητρο που δημιουργεί επιστημονικό ενδιαφέρον για την εκπόνηση της διπλωματικής εργασίας.
- Ουσιαστικά, το θέμα της διπλωματικής εργασίας αποτελεί τον στόχο της έρευνας, ενώ το κίνητρο τεκμηριώνει την πρωτοτυπία και τη συμβολή της.
- Είναι ιδιαίτερα σημαντικό το θέμα και το ερευνητικό κίνητρο να παρουσιάζονται άμεσα και ξεκάθαρα ώστε να γίνεται κατανοητό από τον αναγνώστη αφενός το αντικείμενο της διπλωματικής εργασίας και αφετέρου ο λόγος που παρουσιάζει ενδιαφέρον η περαιτέρω ανάγνωσή της.

Εισαγωγή

- Έστω ότι το θέμα της διπλωματικής εργασίας είναι η διερεύνηση της υποχρεωτικής υιοθέτησης των ΔΠΧΠ από τις ευρωπαϊκές εισηγμένες επιχειρήσεις στην ποιότητα των λογιστικών τους κερδών.
- Ο φοιτητής θα πρέπει να διευκρινίσει άμεσα το θέμα και να καταδείξει το ενδιαφέρον που παρουσιάζει η εμπειρική διερεύνηση αυτού του θέματος, όπως π.χ. να αναφερθεί στο μεγάλο αριθμό εταιρειών που επηρεάζονται, να κάνει μια σύντομη αναφορά σε κάποια στοιχεία που δείχνουν μεγάλες μεταβολές των λογιστικών κερδών μετά την υιοθέτηση, να επισημάνει με συντομία κάποιες προηγούμενες έρευνες που παρουσιάζουν αντικρουόμενα αποτελέσματα, κ.λπ.
- Στη συνέχεια της εισαγωγής, θα πρέπει να αναφέρεται η ερευνητική υπόθεση που διαμορφώθηκε και, εν συντομία, η διαδικασία διερεύνησης που ακολουθήθηκε. Συγκεκριμένα, αναφέρεται η υπόθεση προς διερεύνηση, το δείγμα που συλλέχθηκε, τα βασικά εμπειρικά ευρήματα και τα συμπεράσματα που εξάγονται. Στόχος αυτής της σύντομης παρουσίασης είναι να προϊδεάσει τον αναγνώστη για το γενικότερο περιεχόμενο της διπλωματικής εργασίας.

Εισαγωγή

- Το σημαντικότερο μέρος της εισαγωγής είναι η ανάλυση της πρωτοτυπίας και της συμβολής της ΔΕ στην επιστήμη. Θα πρέπει να αναφερθεί ο λόγος που η συγκεκριμένη ΔΕ εμφανίζει πρωτοτυπία, αλλά και οι λόγοι που μπορεί να φανεί χρήσιμη στις ομάδες που ενδιαφέρονται για το συγκεκριμένο επιστημονικό πεδίο, όπως επαγγελματίες της αγοράς, ρυθμιστικές αρχές, ακαδημαϊκοί, κτλ. Η εισαγωγή ολοκληρώνεται με σύντομη αναφορά στο περιεχόμενο των επόμενων ενοτήτων, όπως π.χ. «στην Ενότητα 2, πραγματοποιείται ανασκόπηση της σχετικής βιβλιογραφίας, στην Ενότητα 3 αναπτύσσονται οι ερευνητικές υποθέσεις», κτλ.
- Σε πρακτικό επίπεδο, η ολοκλήρωση της εισαγωγής, συνήθως, πραγματοποιείται όταν η διπλωματική εργασία βαίνει προς την ολοκλήρωσή της. Ο λόγος είναι ότι μόνο τότε ο φοιτητής έχει αποκρυσταλλώσει τα βασικά αποτελέσματα και συμπεράσματα της έρευνάς του και τη συμβολή της στην επιστήμη.

