

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ  
ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ  
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ  
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ & ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ**

**Π Τ Υ Χ Ι Α Κ Η Ε Ρ Γ Α Σ Ι Α**

**«ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΜΕ ΥΠΟΔΕΙΓΜΑΤΑ PANEL  
(PANEL DATA ANALYSIS)»**

<b>ΙΩΑΝΝΙΔΗΣ ΣΤΑΥΡΟΣ</b>	<b>A.M.: 13224</b>
<b>ΚΑΤΣΑΡΗ ΕΛΕΝΗ</b>	<b>A.M.: 14982</b>
<b>ΚΟΝΤΟΔΗΜΟΣ ΕΥΑΓΓΕΛΟΣ</b>	<b>A.M.: 13880</b>

**ΜΕΣΟΛΟΓΓΙ 2014**

**ΤΕΧΝΟΛΟΓΙΚΟ ΕΚΠΑΙΔΕΥΤΙΚΟ ΙΔΡΥΜΑ  
ΔΥΤΙΚΗΣ ΕΛΛΑΔΑΣ  
ΣΧΟΛΗ ΔΙΟΙΚΗΣΗΣ & ΟΙΚΟΝΟΜΙΑΣ  
ΤΜΗΜΑ ΛΟΓΙΣΤΙΚΗΣ & ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ**

**Π Τ Υ Χ Ι Α Κ Η Ε Ρ Γ Α Σ Ι Α**

**«ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΜΕ ΥΠΟΔΕΙΓΜΑΤΑ PANEL  
(PANEL DATA ANALYSIS)»**

<b>ΙΩΑΝΝΙΔΗΣ ΣΤΑΥΡΟΣ</b>	<b>A.M.: 13224</b>
<b>ΚΑΤΣΑΡΗ ΕΛΕΝΗ</b>	<b>A.M.: 14982</b>
<b>ΚΟΝΤΟΔΗΜΟΣ ΕΥΑΓΓΕΛΟΣ</b>	<b>A.M.: 13880</b>

stayioan@logistiki.teimes.gr

elenkats4@logistiki.teimes.gr

eyagkont@logistiki.teimes.gr

**ΜΕΣΟΛΟΓΓΙ, 2014**

## ΠΡΟΛΟΓΟΣ

Κάθε επιστήμη δε μπορεί να σταθεί χωρίς αναφορά στα εμπειρικά δεδομένα ανεξάρτητα από τη λογική της συνέπεια. Η Οικονομική Επιστήμη είναι μια εμπειρική επιστήμη, σε αντίθεση με τη Μαθηματική Επιστήμη που είναι μια *a priori* επιστήμη. Όμως η έμφαση και η εφαρμογή ποσοτικών μεθόδων στην ανάλυση και την έρευνα των οικονομικών φαινομένων αποτελεί το σπουδαιότερο χαρακτηριστικό των εξελίξεων στην Οικονομική Επιστήμη.

Η οικονομετρία αναφέρεται κυρίως στη ποσοτική πλευρά της Οικονομικής επιστήμης και προσπαθεί να δώσει εμπειρικό περιεχόμενο στις διάφορες αφηρημένες έννοιες και σχέσεις οικονομικής θεωρίας. Οι προβλέψεις αποτελούν αναπόσπαστο τμήμα της σύγχρονης Οικονομικής Επιστήμης αφού το πεδίο εφαρμογής τους επί των ζητημάτων που η τελευταία πραγματεύεται είναι αρκετά ευρύ. Άμεσα συνυφασμένος με την έννοια των προβλέψεων είναι ο παράγοντας του τυχαίου σφάλματος ο οποίος διαφοροποιεί την οικονομετρία από τα γνωστικά αντικείμενα των θετικών επιστημών. Η οικονομετρία είναι ο κλάδος της οικονομικής επιστήμης που σχετίζεται με την εξαγωγή ποσοτικών σχέσεων μεταξύ των μεταβλητών που συνθέτουν ένα οικονομικό φαινόμενο.

Για να γίνει, όμως, κατανοητός ο ρόλος της, θα πρέπει προηγουμένως να γίνει η απαραίτητη εισαγωγή στις θεμελιώδεις έννοιες που αυτή μετέρχεται. Μια από τις πιο θεμελιωδέστερες έννοιες της οικονομετρίας είναι εκείνη του οικονομετρικού υποδείγματος – άμεσα συνυφασμένη με την έννοια του τυχαίου σφάλματος – που εκτενής ανάλυση του περιεχομένου της επιχειρείται κατά την επόμενη παράγραφο του κεφαλαίου αυτού.

Για να καταστεί δυνατή η ανάλυση παλινδρόμησης ενός υποδείγματος, οφείλουν να πληρούνται ορισμένες υποθέσεις για τις οποίες εκτενής αναφορά πρόκειται να γίνει στην τρίτη παράγραφο. Μη εκπλήρωση κάποιας ή κάποιων εκ των υποθέσεων αυτών οδηγεί σε πλειάδα ανεπιθύμητων φαινομένων με σημαντικότερα εκείνα της

αυτοσυσχέτισης, της ετεροσκεδαστικότητας, της πολυσυγγραμμικότητας και της παραβίασης της κανονικότητας.

## ΕΙΣΑΓΩΓΗ

Θεωρία και πράξη είναι δυο αντιπαραβαλλόμενες έννοιες στη ζωή μας. Η χρήση των υπολογιστών ανέδειξε έντονα αυτό το πρότυπο στην επιστήμη. Στις θετικές επιστήμες, και όχι μόνο, αντικειμενικός στόχος είναι αυτές οι δυο έννοιες να έρθουν όσο πιο κοντά είναι δυνατόν. Εργαλεία για την εργασία αυτή είναι κυρίως η προσομοίωση και η πρόβλεψη. Συνδεδειγμένος κρίκος της θεωρίας και της πράξης είναι η δυνατότητα των μαθηματικών εξισώσεων της θεωρίας να συμπίπτουν με τις παρατηρήσεις προσομοίωσης της πράξης. Με άλλα λόγια, η δυνατότητα κατανόησης του παρελθόντος να μας οδηγήσει στην πρόβλεψη του μέλλοντος.

Επιπλέον, βασικός στόχος σχεδόν κάθε ανθρώπινης δραστηριότητας είναι η διενέργεια προβλέψεων, δηλαδή η χρησιμοποίηση ενός εκτιμώμενου μοντέλου για την πρόβλεψη των μελλοντικών τιμών των μεγεθών. Η πρόβλεψη είναι απαραίτητη γιατί αποτελεί το βασικό εργαλείο για κάθε μελλοντική εξέλιξη και απόφαση.

Παράλληλα οι προβλέψεις που γίνονται πρέπει να είναι έγκυρες και όσο πιο κοντά στο μέλλον που θα προκύψει. Η ποιότητα της πρόβλεψης εξαρτάται κυρίως από τον τρόπο που συλλέχτηκαν και αναλύθηκαν οι πληροφορίες. Έτσι, η ανάγκη για έγκυρες, άρα και πρακτικά χρήσιμες προβλέψεις οδήγησε στην ανάπτυξη πολλών τεχνικών και μεθόδων συλλογής και επεξεργασίας διαφόρων πρωτογενών πληροφοριών.

Οι μέθοδοι προβλέψεων μπορεί να είναι ποιοτικές ή ποσοτικές. Με τις ποιοτικές μεθόδους ασχολούνται επιστήμες όπως η φιλοσοφία, θεολογία, οικονομία, πολιτικές επιστήμες κ.α., οι οποίες για την ανάλυση των δεδομένων χρησιμοποιούν την ‘ανθρώπινη’ λογική και τον τρόπο με τον οποίο οι διάφορες σχολές ανάλυσης αντιλαμβάνονται τον φυσικό κόσμο. Με τον όρο ‘ανθρώπινη’ λογική εννοείται κάθε απόφαση που λαμβάνει κανείς μη χρησιμοποιώντας μαθηματικές εξισώσεις. Για παράδειγμα, πρόβλεψη, η καλύτερα εκτίμηση, για την εξέλιξη μιας γλώσσας, της πολιτικής κατάστασης μιας χώρας, την πορεία της Ευρωπαϊκής Ένωσης και άλλα, είναι

αποτέλεσμα λογικών διεργασιών σε διάφορα φόρουμ σκέψης. Σε όλες τις παραπάνω προβλέψεις δεν χρειάζεται καμία μαθηματική επεξεργασία. Αντίθετα, στις ποσοτικές μεθόδους η ανάλυση των δεδομένων γίνεται με την βοήθεια μαθηματικών εκφράσεων υποστηριζόμενες από υπολογιστές. Μια από αυτές τις ποσοτικές μεθόδους, είναι η ανάλυση χρονολογικών σειρών.

Η οικονομετρία είναι η εφαρμογή των μαθηματικών και της στατιστικής για την ανάλυση ποσοτικών και ποιοτικών οικονομικών δεδομένων, με σκοπό να εκτιμηθούν διάφορες οικονομικές σχέσεις, να ελεγχθεί η ορθότητα της οικονομικής θεωρίας και να γίνουν προβλέψεις ως προς την εξέλιξη των οικονομικών φαινομένων.

Οι δύο βασικοί στόχοι της οικονομετρίας είναι να δώσει εμπειρικό περιεχόμενο στην Οικονομική θεωρία και να την υποβάλει σε έλεγχο για πιθανόν λανθασμένες προβλέψεις. Για παράδειγμα η Οικονομική θεωρία μπορεί να προβλέπει ότι μια κάποια καμπύλη ζήτησης οφείλει να έχει αρνητική κλίση, οι Οικονομετρικές προβλέψεις δύνανται να επαληθεύσουν ή όχι την παραπάνω πρόβλεψη και να φωτίσουν το μέγεθος των αποτελεσμάτων της.

Η οικονομετρία μπορεί να αποκαλύψει μόνο συσχετίσεις και όχι αιτιακές σχέσεις. Το ότι αλλαγές σε μία μεταβλητή φαίνεται να συσχετίζονται με αλλαγές σε μία άλλη δεν σημαίνει αναγκαστικά ότι η μία μεταβλητή προκαλεί την άλλη.

## **ΕΠΙΣΗΜΑΝΣΗ**

Οι διαπιστώσεις, τα αποτελέσματα, τα συμπεράσματα και οι πιθανές προτάσεις της παρούσας πτυχιακής εργασίας –εκτός των αναφορών που σημαίνονται ως λήμματα– αποτελούν προσωπικές θεωρητικές ή εμπειρικές διαπιστώσεις του σπουδαστή (σπουδάστριάς ) ή της ομάδας των σπουδαστών που την επιμελήθηκαν και δεν απηχούν κατ' ανάγκη τη γνώμη του εισηγητή εκπαιδευτικού, του Εκπαιδευτικού Προσωπικού του Τμήματος Λογιστικής ή του Α.Τ.Ε.Ι. Μεσολογγίου.

## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ

ΠΡΟΛΟΓΟΣ.....	3
ΕΙΣΑΓΩΓΗ .....	5
ΕΠΙΣΗΜΑΝΣΗ .....	7
ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ.....	8
<b>ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ΣΧΗΜΑΤΩΝ .....</b>	<b>11</b>
<b>ΠΕΡΙΛΗΨΗ.....</b>	<b>13</b>
<b>ΚΕΦΑΛΑΙΟ 1 «ΟΙΚΟΝΟΜΕΤΡΙΚΗ ΑΝΑΛΥΣΗ-ΟΙΚΟΝΟΜΕΤΡΙΚΑ ΜΟΝΤΕΛΑ».....</b>	<b>15</b>
1.1. ΕΝΝΟΙΑ ΚΑΙ ΣΚΟΠΟΣ ΤΗΣ ΟΙΚΟΝΟΜΕΤΡΙΑΣ.....	15
1.1.1. ΟΙΚΟΝΟΜΙΚΟ και ΟΙΚΟΝΟΜΕΤΡΙΚΟ ΥΠΟΔΕΙΓΜΑ .....	16
1.1.2. ΤΥΠΟΙ ΔΕΔΟΜΕΝΩΝ.....	18
1.2. ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΕΙΣ ΠΑΡΑΜΕΤΡΩΝ .....	19
1.3. ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗΣ .....	21
1.4. ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ.....	27
1.5. ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ στο ΑΠΛΟ ΓΡΑΜΜΙΚΟ ΜΟΝΤΕΛΟ.....	32
1.5.1. ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑ .....	35
1.5.2. ΑΥΤΟΣΥΣΧΕΤΙΣΗ.....	40
1.5.3. ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ .....	42
1.5.4. ΣΦΑΛΜΑ ΕΞΕΙΔΙΚΕΥΣΗΣ .....	44
1.5.5. ΟΛΟΚΛΡΩΣΗ - ΣΥΝΟΛΟΚΛΗΡΩΣΗ.....	46
1.5. ΠΡΟΒΛΕΨΕΙΣ .....	47



<b>ΚΕΦΑΛΑΙΟ 2 «ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ – ΑΝΑΛΥΣΗ ΔΙΑΣΤΡΩΜΑΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ» .....</b>	<b>49</b>
2.1. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ – ΧΡΗΣΙΜΟΤΗΤΑ - ΧΡΟΝΟΣΕΙΡΩΝ .....	49
2.2. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ & ΜΕΤΡΑ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ.....	52
2.3. ΣΤΑΣΙΜΕΣ – ΜΗ ΣΤΑΣΙΜΕΣ ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ.....	54
2.5. ΑΥΤΟΠΑΛΙΝΔΡΟΜΑ ΥΠΟΔΕΙΓΜΑΤΑ .....	58
2.6. ΥΠΟΔΕΙΓΜΑΤΑ ΚΙΝΗΤΟΥ ΜΕΣΟΥ.....	62
2.7. ΤΑΣΗ ΚΑΙ ΕΠΟΧΙΚΟΤΗΤΑ.....	64
2.8. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ – ΧΡΗΣΙΜΟΤΗΤΑ – ΔΙΑΣΤΡΩΜΑΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ 65	
2.9. ΠΕΔΙΟ ΕΦΑΡΜΟΓΗΣ.....	67
2.10. ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ .....	69
2.11. ΔΙΑΣΤΡΩΜΑΤΙΚΑ ΣΤΟΙΧΕΙΑ vs ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ .....	70
<b>ΚΕΦΑΛΑΙΟ 3: «ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΠΑΝΕΛ» .....</b>	<b>71</b>
3.1. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ - ΧΡΗΣΙΜΟΤΗΤΑ .....	72
3.1.1. ΠΛΕΟΝΕΚΤΗΜΑΤΑ & ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ PANEL 73	
3.2. ΓΕΝΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΟΣ ΣΕ PANEL DATA.....	78
3.3. ΜΟΝΤΕΛΟ ΜΗΔΕΝΙΚΩΝ ΕΠΙΔΡΑΣΕΩΝ .....	83
3.4. ΜΟΝΤΕΛΟ ΣΤΑΘΕΡΩΝ ΕΠΙΔΡΑΣΕΩΝ .....	84
3.5. ΜΟΝΤΕΛΟ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ.....	87
3.6. ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΣΤΑΘΕΡΩΝ ΚΑΙ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ.....	89

3.7. ΕΛΕΓΧΟΣ HAUSMAN .....	90
<b>ΚΕΦΑΛΑΙΟ 4 «ΕΦΑΡΜΟΓΕΣ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ ΠΑΝΕΛ – E-VIEWS»</b> .....	92
4.1. ΕΦΑΡΜΟΓΗ ΣΤΗ ΝΑΥΤΙΛΙΑ - ΠΑΡΑΔΕΙΓΜΑ .....	92
4.2. ΕΦΑΡΜΟΓΗ ΣΤΟ ΧΡΗΜΑΤΙΣΤΗΡΙΟ (ΕΚΤΙΜΗΣΗ ΜΕΤΟΧΩΝ) - ΠΑΡΑΔΕΙΓΜΑ .....	101
4.3. ΠΑΡΑΔΕΙΓΜΑ - ΑΣΚΗΣΕΙΣ .....	105
<b>4.4. ΤΟ ΟΙΚΟΝΟΜΕΤΡΙΚΟ ΠΑΚΕΤΟ E-VIEWS</b> .....	119
<b>ΣΥΜΠΕΡΑΣΜΑΤΑ</b> .....	121
<b>ΒΙΒΛΙΟΓΡΑΦΙΑ</b> .....	122
<b>ΗΛΕΚΤΡΟΝΙΚΕΣ ΠΗΓΕΣ</b> .....	123

## ΠΙΝΑΚΑΣ ΠΕΡΙΕΧΟΜΕΝΩΝ ΣΧΗΜΑΤΩΝ

Σχήμα 1: Οικονομετρική Ανάλυση.....	21
Σχήμα 2: Ομοσκεδαστικότητα vs Ετεροσκεδαστικότητα .....	23
Σχήμα 3: Συνδιακύμανση Σφάλματος .....	24
Σχήμα 4: Υποθέσεις απλής παλινδρόμησης .....	25
Σχήμα 5: Απλή παλινδρόμηση – Κατάλοιπα - Κλίση .....	30
Σχήμα 6: Αποδοχή – Απόρριψη της Μηδενικής Υπόθεσης $H_0$ .....	33
Σχήμα 7: Ετεροσκεδαστικότητα .....	39
Σχήμα 8: Αυτοσυσχέτιση .....	41
Σχήμα 9: Χρονολογικές Σειρές.....	52
Σχήμα 10: Στασιμότητα ως προς τη διακύμανση λη ως προς το μέσο .....	57
Σχήμα 11: Αυτοσυσχέτιση 1 <sup>ης</sup> τάξης .....	61
Σχήμα 12: Αυτοπαλίνδρομο υπόδειγμα AR(1).....	62
Σχήμα 13: Panel Data - Παλινδρόμηση .....	80
Σχήμα 14: Panel Data - Σφάλματα.....	81
Σχήμα 15: Σφάλμα Επιλεκτικότητας .....	83
Σχήμα 16: Περιγραφική Ανάλυση Κινδύνου.....	97
Σχήμα 17: Μεταβλητότητα Χρονοσειράς.....	98
Σχήμα 18: Μεταβλητότητα FFA.....	99
Σχήμα 19: Μεταβλητότητα FFA ( $\alpha$ ).....	100

Σχήμα 20: Ημερήσιες τιμές κλεισίματος του δείκτη DowJones Utility index .....	102
Σχήμα 21: Ημερήσιες τιμές κλεισίματος του δείκτη DowJones Utility index(log).....	103
Σχήμα 22: Μηνιαίες τιμές αεροπορικών επιβατών.....	104
Σχήμα 23: Μηνιαίες τιμές αεροπορικών επιβατών - Μετασχηματισμός .....	104

## ΠΕΡΙΛΗΨΗ

Η παρούσα πτυχιακή εργασία εκπονείται στα πλαίσια ολοκλήρωσης των προπτυχιακών σπουδών μου. Για την υλοποίηση αυτής οφείλω να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου για την καθοδήγηση και τη στήριξή του, τόσο κατά τη διάρκεια των σπουδών μου όσο και κατά τη διάρκεια εκπόνησης αυτής.

Το θέμα ανάλυσης της εργασίας είναι τα δεδομένα πάνελ και η χρησιμότητα αυτών. Προκειμένου όμως να περιγράψουμε την ανάλυση αυτών, αρχικά θα περιγράψουμε κάποιες εισαγωγικές έννοιες και στη συνέχεια θα δώσουμε και κάποια σχετικά παραδείγματα. Πιο συγκεκριμένα, η δομή και η διάρθρωση της παρούσας πτυχιακής έχει ως εξής:

Το παρόν πόνημα αποτελείται από τέσσερα κεφάλαια. Αναλυτικότερα, στο πρώτο κεφάλαιο αναλύεται η έννοια της οικονομετρικής και οικονομικής ανάλυσης. Οι τύποι δεδομένων αλλά και οι μέθοδοι εκτίμησης αυτών είναι μέρος αυτών των κεφαλαίων. Η ανάλυση συσχέτισης, η ανάλυση παλινδρόμησης και οι γενικότεροι μέθοδοι εκτίμησης αυτών των υποδειγμάτων αποτελούν μέρος αυτού του κεφαλαίου. Οι βασικές υποθέσεις που διέπουν το γραμμικό μοντέλο αλλά και οι έλεγχοι που αφορούν αυτές ολοκληρώνουν το κεφάλαιο αυτό.

Το επόμενο κεφάλαιο αφορά δυο μορφές ανάλυσης δεδομένων, τις χρονολογικές σειρές και τα διαστρωματικά δεδομένα. Οι στόχοι και οι σκοποί της κάθε ανάλυσης, τα επιμέρους μοντέλα – υποδείγματα αλλά και το πεδίο εφαρμογής αυτών αποτελούν το περιεχόμενο του δεύτερου κεφαλαίου.

Το τρίτο κεφάλαιο ασχολείται με την ανάλυση δεδομένων πάνελ. Υπογραμμίζονται τα πλεονεκτήματα και τα μειονεκτήματα της ανάλυσης αυτής καθώς και τα κύρια χαρακτηριστικά αυτής της ανάλυσης. Στο κεφάλαιο αυτό αναλύονται επίσης μοντέλα εκτίμησης δεδομένων πάνελ, το μοντέλο μηδενικών επιδράσεων, το μοντέλο σταθερών επιδράσεων και το μοντέλο τυχαίων επιδράσεων. Ο έλεγχος Hausman αλλά και τα κριτήρια επιλογής μεταξύ των μοντέλων ολοκληρώνουν το κεφάλαιο αυτό.

Το τέταρτο και τελευταίο κεφάλαιο της εργασίας αυτής αφορά την εφαρμογή των δεδομένων πάνελ σε πραγματικά δεδομένα, στη ναυτιλία, τη ζήτηση και το χρηματιστήριο. Το οικονομετρικό πακέτο EVIEWS και η χρησιμότητα αυτού ολοκληρώνουν το συγκεκριμένο κεφάλαιο.

Το παρόν πόνημα ολοκληρώνεται με την εξαγωγή συμπερασμάτων αλλά και την παράθεση των βιβλιογραφικών πηγών από όπου αντλήθηκαν οι απαραίτητες πληροφορίες – γνώσεις για τη συγγραφή αυτής.

# ΚΕΦΑΛΑΙΟ 1 «ΟΙΚΟΝΟΜΕΤΡΙΚΗ ΑΝΑΛΥΣΗ- ΟΙΚΟΝΟΜΕΤΡΙΚΑ ΜΟΝΤΕΛΑ»

Στο παρόν πρώτο κεφάλαιο θα αναφερθούμε στις βασικές έννοιες της οικονομετρίας και της οικονομετρικής ανάλυσης. Θα εντοπίσουμε το σκοπό αυτής αλλά και τα υποδείγματα που χρησιμοποιούνται για την εκτίμηση παραμέτρων σε οικονομικό και οικονομετρικό υπόδειγμα. Επιπλέον θα αναφερθούμε στις μεθόδους εκτίμησης ενός μοντέλου αλλά και στις υποθέσεις που διέπουν εαυτό. Προκειμένου να διαπιστώσουμε αν όντως ισχύουν οι εν λόγω υποθέσεις θα πραγματοποιήσουμε αντίστοιχους ελέγχους τους οποίους και αναλύουμε στις ενότητες του παρόντος κεφαλαίου.

## 1.1. ΕΝΝΟΙΑ ΚΑΙ ΣΚΟΠΟΣ ΤΗΣ ΟΙΚΟΝΟΜΕΤΡΙΑΣ

Η οικονομετρία είναι η εφαρμογή των μαθηματικών και της στατιστικής για την ανάλυση ποσοτικών και ποιοτικών οικονομικών δεδομένων, με σκοπό να εκτιμηθούν διάφορες οικονομικές σχέσεις, να ελεγχθεί η ορθότητα της οικονομικής θεωρίας και να γίνουν προβλέψεις ως προς την εξέλιξη των οικονομικών φαινομένων (Βάμβουκας 2007).

Ο ορισμός της οικονομετρίας περιλαμβάνει την εμπειρική εκτίμηση των οικονομικών σχέσεων. Χρησιμοποιώντας δηλαδή την οικονομική θεωρία την στατιστική θεωρία αλλά και τα απαραίτητα σετ δεδομένων από ένα δείγμα ενός πληθυσμού ελέγχει αλλά και μέτρα τις ορισμένες σχέσεις ανάμεσα σε διάφορες οικονομικές μεταβλητές.

Οι σκοποί της οικονομετρίας είναι βασικά τρεις: Πρώτον ο έλεγχος της οικονομικής θεωρίας ως προς την εμπειρική της τεκμηρίωση παράλληλα με την διερεύνηση των οποίων δυνατοτήτων για αναδιατύπωση κάποιων σχέσεων η οποία και καλείται διαρθρωτική ανάλυση (structural analysis), δεύτερον η διατύπωση εναλλακτικών προτάσεων οικονομικής πολιτικής (policy evaluation) και τρίτον η διενέργεια προβλέψεων όσον αφορά την εξέλιξη των διάφορων τιμών ορισμένων

οικονομικών μεγεθών σε περιοχές του δείγματος εκτίμησης (forecasting).

### **1.1.1. ΟΙΚΟΝΟΜΙΚΟ και ΟΙΚΟΝΟΜΕΤΡΙΚΟ ΥΠΟΔΕΙΓΜΑ**

Οι οικονομικές θεωρίες αντικατοπτρίζονται από τα οικονομικά υποδείγματα τα οποία εμπεριέχουν την οικονομική θεωρία σε αλγεβρική μορφή. Ο ακριβής προσδιορισμός των συντελεστών της αλγεβρικής διατύπωσης είναι ο στόχος της οικονομετρίας.

Τα οικονομικά υποδείγματα περιγράφουν τη συμπεριφορά επιχειρήσεων - καταναλωτών και την αλληλεπίδρασή τους στις διάφορες αγορές. Τα οικονομικά υποδείγματα:

- Αφαιρούν τις πολυπλοκότητες του πραγματικού κόσμου.
- Εστιάζουν στα ουσιώδη χαρακτηριστικά των οικονομικών δραστηριοτήτων.
- Επιτρέπουν τη μελέτη των οικονομικών φαινομένων σε απλουστευμένο πλαίσιο.

Τα γενικά χαρακτηριστικά των οικονομικών υποδειγμάτων μπορούν να συνοψισθούν στα ακόλουθα:

#### **Η Υπόθεση *ceteris paribus*:**

- Μελετάμε την επίδραση ορισμένων παραγόντων στην υπό εξέταση μεταβλητή (π.χ. επίδραση μισθών αγροτών στις τιμές των σιτηρών) υποθέτοντας ότι οι υπόλοιποι παράγοντες παραμένουν αμετάβλητοι.
- Προβληματική εμπειρική επαλήθευση της θεωρίας από πραγματικά δεδομένα.
- Ανάπτυξη στατιστικών μεθόδων για τον περιορισμό των “εξωτερικών επιδράσεων”.

#### **Η Υπόθεση της Βελτιστοποίησης**



- Οι δρώντες οικονομικοί παράγοντες (καταναλωτές, επιχειρήσεις, κυβέρνηση κ.λπ.) επιδιώκουν ορθολογικά την επίτευξη κάποιου στόχου.
- Παράδειγμα: Οι επιχειρήσεις επιδιώκουν τη μεγιστοποίηση των κερδών τους.
- Χρήση μαθηματικών μεθόδων βελτιστοποίησης για να προβλεφθεί η συμπεριφορά των δρώντων παραγόντων.

### **Διάκριση μεταξύ Θετικού και Κανονιστικού**

- Θετική οικονομική θεωρία: εξηγεί τα παρατηρούμενα οικονομικά φαινόμενα/ διασαφηνίζει πώς κατανέμονται πραγματικά οι πόροι στην οικονομία (“*τι γίνεται*”).
- Κανονιστική οικονομική θεωρία: διασαφηνίζει πώς πρέπει να κατανέμονται οι πόροι στην οικονομία (“*τι πρέπει να γίνει*”).

Το πρώτο βήμα στην πρακτική της οικονομετρίας είναι η διατύπωση ενός οικονομετρικού υποδείγματος. Τα οικονομετρικά υποδείγματα διακρίνονται σε γραμμικά και μη γραμμικά, εκτός από μια εξίσωση συχνά μπορούν να αποτυπωθούν και ως σύστημα εξισώσεων. Στο σημείο αυτό να αναφέρουμε πως κάποια υποδείγματα ονομάζονται δυναμικά, πρόκειται για τα υποδείγματα τα οποία εμφανίζουν χρονικές υστερήσεις. Οι χρονικές υστερήσεις προκύπτουν από το γεγονός ότι η μεταβολή ενός προσδιοριστικού παράγοντα χρειάζεται κάποιο χρονικό διάστημα για να επιδράσει.

Ένα οικονομετρικό υπόδειγμα αποτελείται από τα παρακάτω δύο μέρη:

- **το συστηματικό μέρος**, το οποίο περιλαμβάνει τις οικονομικές μεταβλητές και
- **το τυχαίο μέρος**, το οποίο περιλαμβάνει το τυχαίο σφάλμα.

Κατά το κυρίως στάδιο της οικονομικής ανάλυσης, εκείνο, δηλαδή, της ανάλυσης παλινδρόμησης, πραγματοποιείται η εκτίμηση ενός οικονομετρικού υποδείγματος, δηλαδή προσδιορισμός των τιμών των παραμέτρων του, ενώ, επιπροσθέτως, διενεργούνται διάφοροι στατιστικοί έλεγχοι με σκοπό να διερευνηθεί η ορθότητα της ως άνω εκτίμησης. Ένα οικονομετρικό υπόδειγμα καλείται γραμμικό όταν παρουσιάζει γραμμικότητα ως προς τις παραμέτρους του.

Τα υποδείγματα που χρησιμοποιούνται στην οικονομική θεωρία ονομάζονται ντετερμινιστικά ενώ τα οικονομετρικά υποδείγματα ονομάζονται στοχαστικά.

Τα οικονομετρικά υποδείγματα όπως ήδη αναφέρθηκε διακρίνονται σε γραμμικά (linear) και μη γραμμικά (non – linear). Μερικές από τις πιο γνωστές μορφές αυτών είναι οι ακόλουθες:

$$Y = \beta_0 + \beta_1 X + \epsilon$$

$$Y = \beta_0 + \beta_1 X + \beta_2 XZ + \epsilon$$

$$\log Y = \beta_0 + \beta_1 \log X + \epsilon$$

$$Y = \beta_0 + \beta_1 \log X + \epsilon$$

$$\log Y = \beta_0 + \beta_1 \log X + \epsilon$$

$$Y = \beta_0 + \beta_1 \frac{1}{X} + \epsilon$$

$$Y = \beta_0 X^{\beta_1} + \beta_2 X^{\beta_2} + \epsilon$$

$$\log\left(\frac{Y}{1-Y}\right) = \beta_0 + \beta_1 X + \epsilon$$

$$Y = \beta_0 e^{\beta_1 X}$$

### 1.1.2. ΤΥΠΟΙ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα είναι μία σειρά από παρατηρήσεις που συλλέγονται σε τακτικά χρονικά διαστήματα. Τέτοια δεδομένα είναι πολύ συχνά:

- Στα οικονομικά (ανεργία ,πληθωρισμός)
- Περιβαλλοντικά οικονομικά (βροχόπτωση, πληθυσμοί σπάνιων ειδών)
- Χρηματοοικονομικά (τιμές μετοχών, ισοτιμίες)
- Ιατρική (εγκεφαλική λειτουργία ανά sec).

Ορισμένοι από τους πιο συχνά συναντώμενους τύπους δεδομένων είναι οι εξής:

- **Στοιχεία χρονολογικών σειρών:** Μια μεταβλητή αποκαλείται χρονολογική σειρά όταν λαμβάνει τιμές κατά τη διάρκεια ορισμένης χρονικής περιόδου. Τα δεδομένα χρονοσειρών έχουν μία διαφορετική παρατήρηση για κάθε χρονική περίοδο – π.χ. τιμές μετοχών. Αφού δεν αποτελούν ένα τυχαίο δείγμα, έχουμε να αντιμετωπίσουμε ένα διαφορετικό πρόβλημα. Σημαντικά θέματα είναι η τάση και η εποχικότητα.

- **Διαστρωματικά Στοιχεία:** Πρόκειται για στοιχεία – στατιστικά δεδομένα τα οποία αφορούν μια ή περισσότερες μεταβλητές σε κάποια συγκεκριμένη χρονολογική στιγμή. Κάθε παρατήρηση είναι ένα νέο άτομο, εταιρία κλπ. με πληροφορίες για κάθε δεδομένη χρονική στιγμή. Εάν τα δεδομένα δεν αποτελούν τυχαίο δείγμα , τότε παρουσιάζεται πρόβλημα στην επιλογή δείγματος.
- **Ομαδοποιημένα Στοιχεία:** Τα ομαδοποιημένα στοιχεία αποτελούνται από δεδομένα τα οποία προκύπτουν από συνδυασμό χρονολογικών σειρών και διαστρωματικών στοιχείων.
- **Στοιχεία Πάνελ:** Τα δεδομένα πάνελ αποτελούν ειδική περίπτωση των ομαδοποιημένων στοιχείων και αξιοποιούνται εκτενώς στη σύγχρονη οικονομετρία (είναι αυτά με τα οποία θα ασχοληθούμε στη συνέχεια της εν λόγω εργασίας). Αποτελούν ένα σύνολο δεδομένων όπου ένα διαστρωματικό δείγμα απεικονίζεται διαχρονικά. Είναι ουσιαστικά συνδυασμός χρονολογικών και διαστρωματικών στοιχείων με κάποια πλεονεκτήματα (περιορίζεται η πολυσυγγραμμικότητα και η ετεροσκεδαστικότητα και επιπλέον αυξάνεται το πλήθος των παρατηρήσεων. Μέσω αυτών μπορούμε να ενώσουμε τυχαία διαστρώματα και να τα επεξεργαστούμε όμοια σαν ένα κανονικό διάστρωμα. Θα χρειαστεί να υπολογίσουμε μόνο τις χρονικές διαφορές. Μπορούμε να παρακολουθήσουμε τις ίδιες τυχαίες παρατηρήσεις με την πάροδο του χρόνου – γνωστές ως ενοποιημένα δεδομένα ή μακροχρόνια δεδομένα.

## 1.2. ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΕΙΣ ΠΑΡΑΜΕΤΡΩΝ

Αναφερόμενοι στην οικονομετρική ανάλυση θα πρέπει να τονίσουμε και τα στάδια αυτής. Το πρώτο στάδιο αφορά την εξειδίκευση του υποδείγματος δηλαδή στον καθορισμό των μεταβλητών που θα το απαρτίζουν , στην καταγραφή αυτών σε εξωγενείς και ενδογενείς καθώς και στην μαθηματική διατύπωση του υποδείγματος. Το δεύτερο

στάδιο αναφέρεται στην κατάλληλη επιλογή των οικονομετρικών τεχνικών για την εκτίμηση των συντελεστών των μεταβλητών μας και ονομάζεται εκτίμηση του υποδείγματος. Τέλος το τρίτο στάδιο αφορά τον έλεγχο του υποδείγματος με την παράλληλη εφαρμογή οικονομικών, στατιστικών αλλά και οικονομετρικών κριτηρίων για το έλεγχο των αποτελεσμάτων της εκτιμήσεως.

Πιο συγκεκριμένα, η διαδικασία της οικονομετρικής ανάλυσης περιλαμβάνει τα εξής στάδια:

- **Εξειδίκευση του υποδείγματος:**
  - Καθορισμός των μεταβλητών που θα περιληφθούν, διαχωρισμός σε ενδογενείς/ εξωγενείς, μαθηματική διατύπωση
  - Η οικονομική θεωρία είναι αυτή η οποία μπορεί να υποδείξει ποιες μεταβλητές είναι σημαντικές ή ίσως σχετικές αλλά δεν καθορίζει την μαθηματική μορφή που συνδέει τις μεταβλητές.
  - Συνήθως η επιλογή της μαθηματικής μορφής της συναρτησιακής σχέσεως είναι συνδυασμός των πληροφοριών από την οικονομική θεωρία και τα πραγματικά δεδομένα
- **Εκτίμηση του υποδείγματος:**
  - Εφαρμογή των κατάλληλων οικονομετρικών μεθόδων για την εκτίμηση των παραμέτρων του υποδείγματος
  - Οι στατιστικές παρατηρήσεις που μπορούν να χρησιμοποιηθούν, όπως ήδη αναφέρθηκε είναι: Χρονολογικές σειρές (time series data): Διαχρονικές παρατηρήσεις για μια σειρά ετών, μηνών κτλ για μια οικονομική μονάδα.
  - Διαστρωματικά στοιχεία (cross-sectional data): Παρατηρήσεις για έναν αριθμό οικονομικών μονάδων για μια χρονική στιγμή.
  - Δυναμικά διαστρωματικά στοιχεία (panel data): Διαχρονικές παρατηρήσεις για μια σειρά οικονομικών μονάδων.
- **Έλεγχος του υποδείγματος:**
  - Αξιολόγηση και έλεγχος των αποτελεσμάτων της εκτιμήσεως.
  - Επιβεβαιώνεται ή αμφισβητείται η οικονομική θεωρία;

Γραφικά όλα τα παραπάνω απεικονίζονται στο ακόλουθο διάγραμμα:



Σχήμα 1: Οικονομετρική Ανάλυση

### 1.3. ΑΝΑΛΥΣΗ ΣΥΣΧΕΤΙΣΗΣ ΚΑΙ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Ο προσδιορισμός της σχέσης ανάμεσα σε δυο ή περισσότερες μεταβλητές ονομάζεται ανάλυση παλινδρόμησης. Πρωταρχική επιδίωξη της ανάλυσης παλινδρόμησης είναι η εκτίμηση ενός οικονομετρικού υποδείγματος, προκειμένου να προσδιοριστεί με ακρίβεια η σχέση μεταξύ της εξαρτημένης και της ανεξάρτητης μεταβλητής (των ανεξάρτητων μεταβλητών).

Το πρόβλημα παλινδρόμησης βασίζεται στην προσπάθεια πρόβλεψης της συμπεριφοράς μιας μεταβλητής (εξαρτημένης), βασισμένη σε μια άλλη (ανεξάρτητη). Όταν αυτή η πρόβλεψη γίνεται σε δύο μόνο τυχαίες μεταβλητές τότε θα μιλάμε για την

απλή παλινδρόμηση, ενώ όταν η πρόβλεψη για την εξαρτημένη μεταβλητή βασίζεται σε περισσότερες από μία μεταβλητές τότε θα ονομάζεται πολλαπλή παλινδρόμηση.

Η ανάλυση παλινδρόμησης δημιουργεί μια στοχαστική σχέση την οποία προσπαθεί να προσεγγίσει. Προκειμένου να κάνει κάτι τέτοιο βασίζεται σε ορισμένες υποθέσεις:

Έστω  $Y_i = \beta_0 + \beta_1 * X_i + e_i$  η στοχαστική σχέση, όπου:

Y: ονομάζεται εξαρτημένη ή ερμηνευόμενη μεταβλητή.

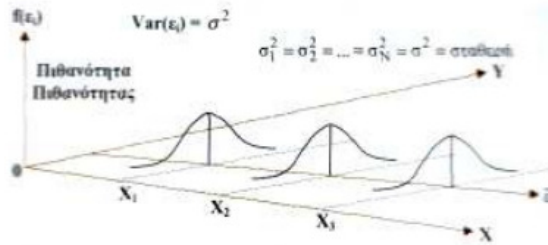
X: είναι η ανεξάρτητη ή ερμηνευτική μεταβλητή και

i: είναι μία αντιπροσωπευτική από τις η παρατηρήσεις του δείγματος.

Η παραπάνω σχέση ονομάζεται εξίσωση γραμμικής παλινδρόμησης και οι υποθέσεις που συνιστούν το κλασικό γραμμικό υπόδειγμα παλινδρόμησης είναι οι εξής:

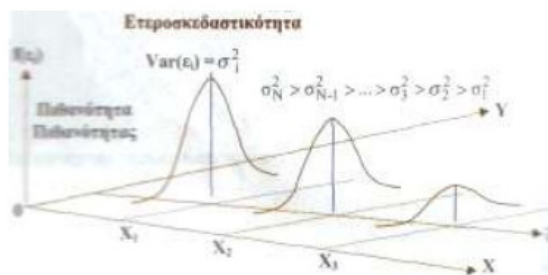
- Η συναρτησιακή μορφή του υποδείγματος είναι γραμμική:  $Y_i = \beta_0 + \beta_1 * X_i + e_i$ , δηλαδή ότι η μαθηματική μορφή η οποία και συνδέει την ανεξάρτητη με την εξαρτημένη μεταβλητή είναι γραμμικής μορφής. Η γραμμικότητα αυτή αναφέρεται στους συντελεστές παλινδρόμησης και όχι στις μεταβλητές του υποδείγματος.
- Ο μέσος του όρου σφάλματος είναι μηδέν:  $E(e_i / X_i) = 0$ , δηλαδή η μεταβλητή  $e_i$  είναι μια τυχαία μεταβλητή η οποία και μπορεί να παίρνει τόσο αρνητικές αλλά και θετικές τιμές αλλά η μέση της τιμή (μαθηματική ελπίδα), υπό τον περιορισμό ότι η τιμή των ανεξάρτητων μεταβλητών είναι δεδομένες, είναι μηδέν. Η σημασία της υπόθεσης αυτής συνίσταται στο γεγονός ότι οι μη εμφανείς παράγοντες οι οποίοι και «υπολογίζονται» στον διαταρακτικό όρο δεν επηρεάζουν συστηματικά την μέση τιμή της εξαρτημένης μεταβλητής
- Η διακύμανση όλων των όρων σφάλματος είναι η ίδια σταθερά  $Var(e_i / X_i) = \sigma^2$ . Η υπόθεση αυτή μας λέει ότι η διασπορά των τιμών της τυχαίας μεταβλητής γύρω από τον μέσο της δεν αλλάζει όταν μεταβάλλεται η τιμή της ανεξάρτητης μεταβλητής  $X_i$ . Όταν η διακύμανση παραμένει σταθερή ο διαταρακτικός όρος χαρακτηρίζεται ομοσκεδαστικός ενώ όταν η διακύμανση δεν είναι σταθερή

ετεροσκεδαστικός.



Ομοσκεδαστικότητα.

(Πηγή: Οικονομετρία  
Α. Ανδρικόπουλου)



Ετεροσκεδαστικότητα

(Πηγή: Οικονομετρία  
Α. Ανδρικόπουλου)

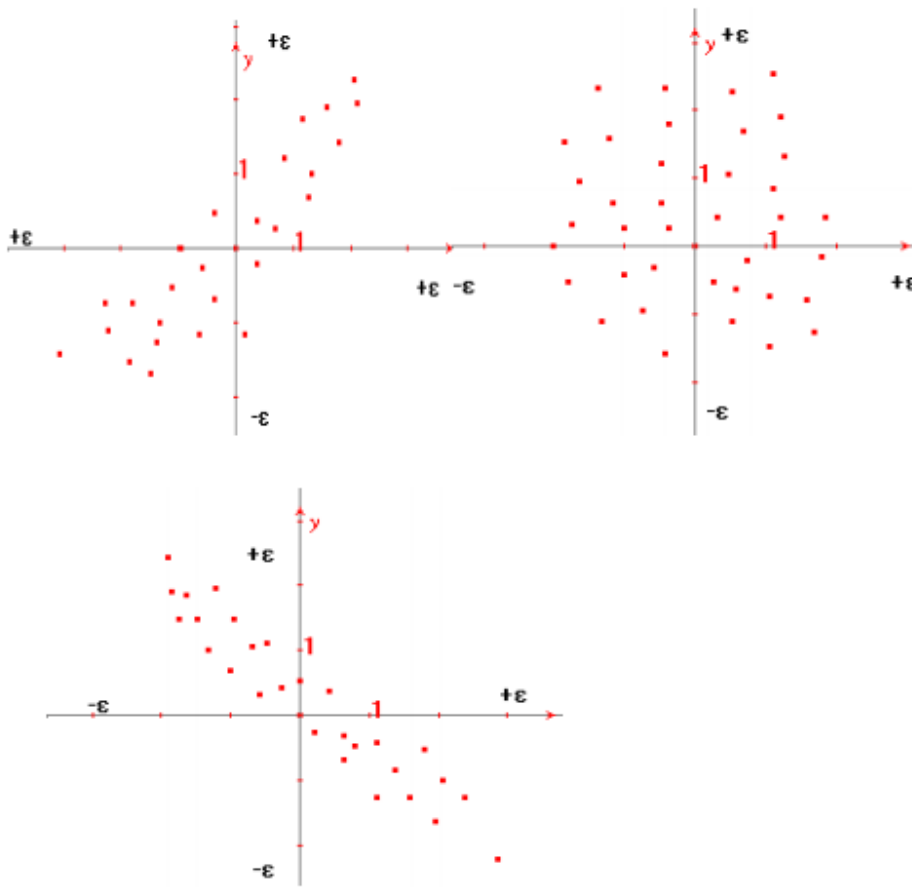
Σχήμα 2: Ομοσκεδαστικότητα vs Ετεροσκεδαστικότητα<sup>1</sup>

- Η συνδιακύμανση μεταξύ των όρων σφάλματος είναι μηδέν:  $Cov(\varepsilon_i, \varepsilon_j) = 0$ .

$$Cov(\varepsilon_i, \varepsilon_j) = E[\varepsilon_i - E(\varepsilon_i)][\varepsilon_j - E(\varepsilon_j)] = E(\varepsilon_i, \varepsilon_j) = 0$$

Δηλαδή η σχέση αυτή μας λέει ότι οι διαταρακτικοί όροι χαρακτηρίζονται από την απουσία της αυτοσυσχέτισης καθώς και ότι για κάθε  $X_i$  οι αποκλίσεις των κάθε τιμών  $Y$  από τις μέσες τιμές δεν μας δίνουν υποδείγματα των κάτωθι μορφών.

<sup>1</sup> Ανδρικόπουλος, Ανδρέας Α. Οικονομετρία : Βασική θεωρία και εφαρμογές / Ανδρέας Α. Ανδρικόπουλου. - 3η έκδ. - Αθήνα : Μπένου Ε., 2003



Σχήμα 3: Συνδιακύμανση Σφάλματος<sup>2</sup>

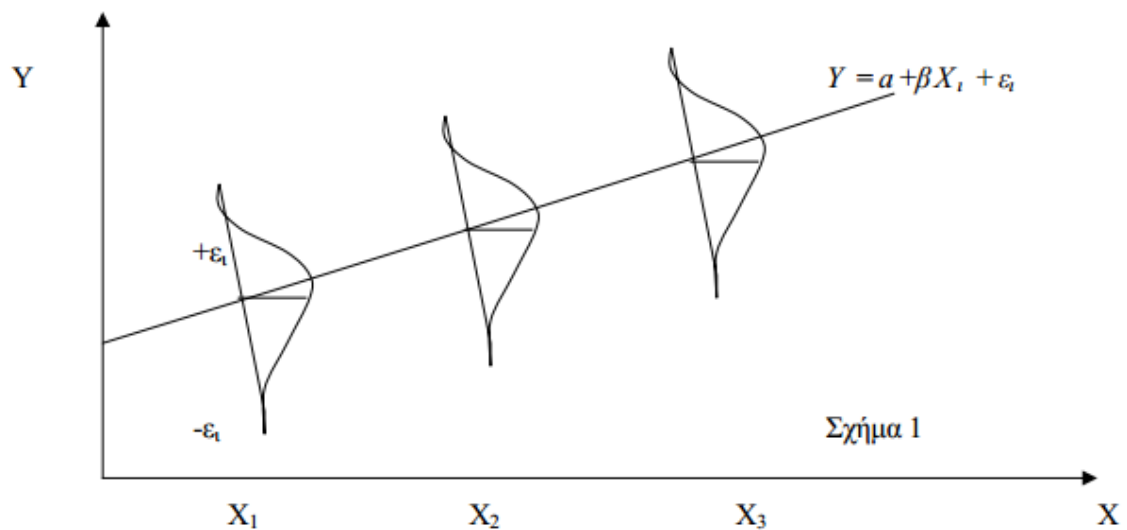
- Η συνδιακύμανση των όρων σφάλματος και των παρατηρήσεων της ανεξάρτητης μεταβλητής είναι πάντα μηδέν:  $\text{Cov}(\varepsilon_i, X_i) = 0$ , για κάθε  $i, j = 1, 2, \dots, n$ . Η υπόθεση αυτή μας τονίζει η ανεξάρτητη μεταβλητή  $X$  δεν είναι στοχαστική και πως οι τιμές παραμένουν σταθερές σε μια επαναληπτική διαδικασία.
- Οι όροι σφάλματος, ανεξάρτητοι μεταξύ τους, ακολουθούν την κανονική κατανομή:  $\varepsilon_i \sim N(0, \sigma^2)$  με μέσο  $\mu = 0$  και διασπορά  $\sigma^2$ . Η τελευταία

<sup>2</sup> Ανδρικόπουλος, Ανδρέας Α. Οικονομετρία : Βασική θεωρία και εφαρμογές / Ανδρέα Α. Ανδρικόπουλου. - 3η έκδ. - Αθήνα : Μπένου Ε., 2003



υπόθεση τίθεται για μικρά δείγματα, όπου μικρά δείγματα στην οικονομετρία θεωρούνται αυτά με αριθμό παρατηρήσεων κάτω από 30 ή κάτω από 20. Η υπόθεση αυτή δεν είναι αναγκαία για μεγάλα δείγματα, αφού βάσει του κεντρικού οριακού θεωρήματος  $\varepsilon_i \sim N(0, \sigma^2)$ . Η τελευταία υπόθεση μας διευκολύνει στην στατιστική επαγωγή και την κατασκευή ελέγχων υποθέσεων σχετικά με την συμπεριφορά των εκτιμητριών.

Οι υποθέσεις αυτές μπορούν να συνοψισθούν στο σχήμα που ακολουθεί:



Σχήμα 4: Υποθέσεις απλής παλινδρόμησης<sup>3</sup>

Η ανάλυση παλινδρόμησης είναι μια στατιστική μέθοδος που προσπαθεί να ερμηνεύσει και να ποσοτικοποιήσει τις μεταβολές μιας μεταβλητής (εξαρτημένη) σε σχέση με τις μεταβολές άλλων μεταβλητών (ανεξάρτητες). Ο ερευνητής με βάση την θεωρία και την εμπειρία επιλέγει:

- Την εξαρτημένη μεταβλητή
- Τις ανεξάρτητες μεταβλητές
- Την μορφή της συνάρτησης
- Και εκτιμά τις παραμέτρους ( $\beta_i$ ).

<sup>3</sup> <http://slideplayer.gr/slide/1900726/>

Οι σχέσεις που αναλύονται μέσω ενός υποδείγματος  $n$ -

Μια από τις μεθόδους που χρησιμοποιούνται για την εκτίμηση της γραμμής παλινδρόμησης είναι η μέθοδος των ελαχίστων τετραγώνων (OLS).

Είναι η μέθοδος που χρησιμοποιείται περισσότερο επειδή

- οι εκτιμητές έχουν πολλές επιθυμητές ιδιότητες
- Είναι εύκολη στην εφαρμογή της.

Ο αριθμός των εκτιμητών είναι ουσιαστικά άπειρος δηλαδή μπορούμε να κατασκευάσουμε και άπειρες γραμμές παλινδρομήσεως. Χρειαζόμαστε ένα κριτήριο. Το κριτήριο αυτό παρουσιάζεται ακολούθως:

Επιλογή των συντελεστών  $\beta_0$  και  $\beta_1$  που ελαχιστοποιούν τα τετράγωνα των αποκλίσεων της ευθείας παλινδρόμησης από τις πραγματικές τιμές.

Σε αυτό το σημείο πρέπει να αναφέρουμε ότι η ακρίβεια της εξίσωσης παλινδρόμησης εξαρτάται από τρεις κύριους παράγοντες.

- Από το μέγεθος του δείγματος. Αφού το  $n$  εμφανίζεται και στον παρανομαστή του κλάσματος όσο μεγαλύτερο είναι το μέγεθος τόσο και μικρότερη η τιμή του standard error άρα τόσο πιο ακριβής η εκτίμηση αυτού.
- Επίσης όσο μεγαλύτερη είναι η απόκλιση του  $X$  τόσο και μικρότερη η τιμή του standard error άρα τόσο πιο ακριβής η εκτίμηση αυτού.
- $\text{Var}(\epsilon)$ . Όσο μεγαλύτερη είναι η απόκλιση των  $\epsilon_i$  στην κατακόρυφη διεύθυνση τόσο το λιγότερο ακριβής είναι η εκτίμηση του standard error of the slope parameter.

Επιπλέον να αναφέρουμε πως στην εκτίμηση παλινδρόμησης σημαντικό ρόλο διαδραματίζει μεταξύ άλλων και ο συντελεστής προσδιορισμού. Η αναλογία της μεταβλητικότητας της εξαρτημένης μεταβλητής που ερμηνεύεται από την παλινδρόμηση ονομάζεται συντελεστής προσδιορισμού (coefficient of determination). Ο συντελεστής προσδιορισμού  $R^2$  μετράει το ποσοστό της μεταβλητικότητας της μεταβλητής  $Y$  η οποία και ερμηνεύεται από την παλινδρόμηση του δείγματος. Ο συντελεστής προσδιορισμού

$R^2$  συμπίπτει με το  $r^2$ , όπου  $r$  ο δειγματικός συντελεστής συσχέτισης στο απλό γραμμικό μοντέλο. Επίσης όταν το  $R^2$  είναι κοντά στην μονάδα τότε μιλάμε ότι το μοντέλο μας είναι καλό στην ερμηνεία της απόκλισης του  $Y$ , υπάρχει τέλεια προσαρμογή της ευθείας παλινδρόμησης (Fit Model).

#### 1.4. ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Σε πολλά πειράματα υπάρχει μία γραμμική σχέση ανάμεσα στα μετρούμενα μεγέθη. Για παράδειγμα, η ταχύτητα ενός σώματος το οποίο εκτελεί ελεύθερη πτώση, μεταβάλλεται γραμμικά με το χρόνο, εφόσον αγνοήσουμε την αντίσταση του αέρα. Τοποθετώντας τα σημεία σε ένα διάγραμμα, βλέπουμε ότι αυτά προσεγγίζουν μία ευθεία γραμμή. Το επόμενο βήμα είναι να βρούμε την κλίση της ευθείας η οποία προσεγγίζει περισσότερο αυτά τα σημεία, και το σημείο στο οποίο αυτή τέμνει τον άξονα  $y$  (τεταγμένη). Σε κάθε περίπτωση, δεν περιμένουμε η ευθεία να διέρχεται από όλα τα σημεία, λόγω της παρουσίας τυχαίων σφαλμάτων. Μπορούμε να βρούμε προσεγγιστικές τιμές τόσο για την κλίση όσο και για την τεταγμένη, εάν σχεδιάσουμε μία ευθεία η οποία να διέρχεται ανάμεσα από τα διεσπαρμένα σημεία. Η ακριβέστερη όμως μέθοδος για να το πετύχουμε αυτό είναι η μέθοδος των ελαχίστων τετραγώνων.

Αναφέραμε ότι γραμμική ευθεία παλινδρόμησης είναι  $a + \beta \cdot x$ , και γνωρίζοντας ότι είναι αδύνατο να βρούμε ένα ακριβώς  $y$ , το οποίο να εισέρχεται από την ευθεία παλινδρόμησης, όταν γνωρίζουμε ένα  $x$ , τότε έχουμε κάποιο προβλεπόμενο  $y$ . Δηλαδή, την ευθεία πρόβλεψης ή παλινδρόμησης ή ελαχίστων τετραγώνων:

$$\hat{y} = \hat{\alpha} + \hat{\beta}x$$

Η εκτίμηση

$$\hat{y} = \hat{\alpha} + \hat{\beta}x$$

της πληθυσμιακής ευθείας παλινδρόμησης  $E(Y / X) = a + \beta \cdot X$ , ονομάζεται ευθεία ελαχίστων τετραγώνων από τη μέθοδο υπολογισμού των παραμέτρων της. Τα

“καπελομένα”  $\alpha$  και  $\beta$  είναι εκτιμητές των  $\alpha$  και  $\beta$ , που επιλέγονται με την μέθοδο των ελαχίστων τετραγώνων. Και επιλέγονται έτσι ώστε το άθροισμά των τετραγώνων των σφαλμάτων να είναι ελάχιστο. Δηλαδή:

$$\text{Ελάχιστο τετράγωνο σφαλμάτων} = \sum_{i=1}^n (Y_i - \alpha - \beta x_i)^2$$

Η γραμμή ελαχίστων τετραγώνων είναι αυτή που ελαχιστοποιεί το άθροισμα των τετραγώνων των σφαλμάτων σε σχέση με τις προσεγγίσεις των  $\alpha$  και  $\beta$ . Η ελαχιστοποίηση της τελευταίας έκφρασης γίνεται σε σχέση με τα  $\alpha$  και  $\beta$ . Μετά από την παραγοντοποίηση αυτής αναφορικά με τα  $\alpha$  και  $\beta$ , αφού τεθούν οι πρώτες παράγωγοι ίσον με το μηδέν τότε έχουμε τις εξισώσεις:

$$\begin{aligned} \alpha n + \beta \sum_{i=1}^n x_i &= \sum Y_i \\ \alpha \sum x_i + \beta \sum_{i=1}^n x_i^2 &= \sum x_i Y_i \end{aligned}$$

Επίσης το άθροισμα των τετραγώνων, καθώς και τα γινόμενά τους υπολογίζονται από τους παρακάτω τύπους:

$$\begin{aligned} SS_x &= \sum (x - \hat{x})^2 = \sum x^2 - \frac{(\sum x)^2}{n} \\ SS_y &= \sum (y - \hat{y})^2 = \sum y^2 - \frac{(\sum y)^2}{n} \\ SS_{xy} &= \sum (x - \hat{x})(y - \hat{y}) = \sum xy - \frac{\sum x \sum y}{n} \end{aligned}$$

Και καταλήγουμε στους εξής τύπους:

$$\hat{\beta} = \frac{\sum_{i=1}^n x_i Y_i - n\bar{x}\bar{Y}}{\sum_{i=1}^n x_i^2 - n(\bar{x})^2}$$

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{x}$$

Χρησιμοποιώντας τη μέθοδο ελαχίστων τετραγώνων σκοπός μας είναι να εκτιμήσουμε τις παραμέτρους του υποδείγματος της παλινδρόμησης, δηλαδή τους συντελεστές  $\beta_0$  και  $\beta_1$  κατά τέτοιο τρόπο, ώστε η ευθεία γραμμή που θα προκύψει να περιγράφει κατά τον καλύτερο δυνατό τρόπο τη σχέση μεταξύ των μεταβλητών  $X$  και  $Y$ . Η γραμμή της παλινδρόμησης πρέπει να περνάει κοντά από τα σημεία που αντιστοιχούν στα ζεύγη των παρατηρήσεων  $(X_i, Y_i)$ , έτσι ώστε να ελαχιστοποιούνται τα σφάλματα της πρόβλεψης. Για να γίνει κατανομή η μέθοδος, πρέπει πρώτα να εισάγουμε μερικούς συμβολισμούς.

Έχει επικρατήσει στη διεθνή βιβλιογραφία να συμβολίζουμε με μικρούς ελληνικούς χαρακτήρες τις τιμές των παραμέτρων του πληθυσμού και με λατινικούς χαρακτήρες τις εκτιμήσεις τους από τα δεδομένα του δείγματος. Τον ίδιο συμβολισμό θα χρησιμοποιήσουμε και εδώ. Οι εκτιμήσεις από τα δεδομένα του δείγματος των συντελεστών παλινδρόμησης του πληθυσμού  $\beta_0$  και  $\beta_1$  είναι οι συντελεστές  $b_0$  και  $b_1$  αντίστοιχα. Μόλις οι εκτιμήσεις αυτές γίνουν γνωστές, θα είμαστε σε θέση να προβλέπουμε τις τιμές της  $Y$  με την εξίσωση παλινδρόμησης:

$$\hat{Y} = b_0 + b_1 X$$

Δηλαδή, η  $\hat{Y}$  είναι η εκτίμηση της  $E(Y)$ . Έτσι, κατά αναλογία με την εξίσωση που αναφέρεται στη γραμμή παλινδρόμησης του πληθυσμού, οι αποκλίσεις μεταξύ των πραγματικών τιμών της  $Y$  και των τιμών  $\hat{Y}$  με  $e$  (όπου  $e$  τα κατάλοιπα που προκύπτουν από τις προσεγγιστικές τιμές της γραμμής παλινδρόμησης  $b_0 + b_1 * X$  στο σύνολο των  $n$  σημείων), δηλαδή:

$$e_i = Y_i - \hat{Y}_i$$

ή

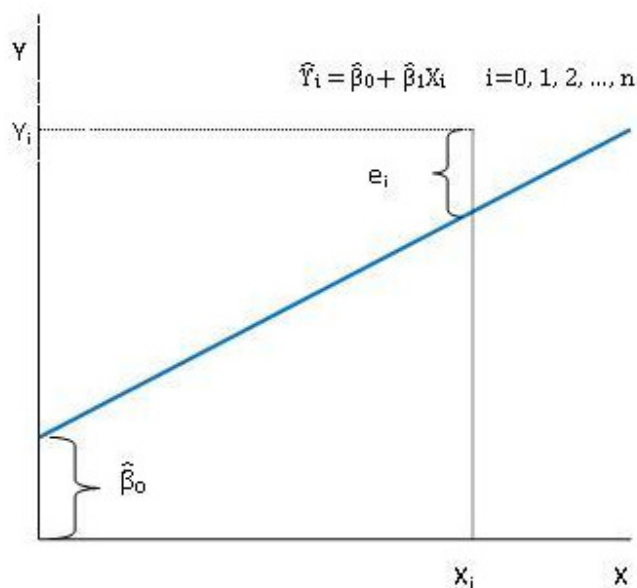
$$e_i = Y_i - (b_0 + b_1 X_i) \quad , \text{ για } i = 1, \dots, n$$

Επομένως, αναζητούμε εκείνες τις τιμές των  $b_0$  και  $b_1$  που θα ελαχιστοποιούν τις αποκλίσεις ( κατάλοιπα ή σφάλματα )  $e_i$ .

Επειδή τα σφάλματα έχουν και θετικό και αρνητικό πρόσημο, θα προσπαθήσουμε να ελαχιστοποιήσουμε τα τετράγωνα τους και μάλιστα το άθροισμά τους. Έτσι λοιπόν προκύπτει και η ονομασία της μεθόδου των ελαχίστων τετραγώνων. Το άθροισμα των τετραγώνων των αποκλίσεων για τα ζεύγη  $n$  ζεύγη των παρατηρήσεων ισούται με :

$$\sum_{i=1}^n (Y_i - \hat{Y}_i)^2 = \sum [Y_i - (b_0 + b_1 X_i)]^2 .$$

Έστω ότι για την παραπάνω εξίσωση έχουν εκτιμηθεί οι συντελεστές παλινδρόμησης  $\beta_0 = \hat{\beta}_0$  και  $\beta_1 = \hat{\beta}_1$ . Ο στόχος μας είναι οι τιμές αυτές να αποκλίνουν όσο το δυνατόν λιγότερο. Έτσι προκύπτει η γραμμή παλινδρόμησης:



Σχήμα 5: Απλή παλινδρόμηση – Κατάλοιπα - Κλίση<sup>4</sup>

<sup>4</sup> <http://slideplayer.gr/slide/1900726/>

Το Θεώρημα Gauss - Markov ισχύει μόνο για γραμμικά συστήματα παλινδρόμησης και λέει ότι δεδομένων των υποθέσεων, οι συντελεστές  $\widehat{\beta}_0$  και  $\widehat{\beta}_1$  είναι οι πιο αποτελεσματικοί από όλους τους πιθανούς αμερόληπτους εκτιμητές των  $\beta_0$  και  $\beta_1$ , καθώς έχουν την μικρότερη διακύμανση μεταξύ των οποιονδήποτε αμερόληπτων εκτιμητών. Αν το θεώρημα δεν ισχύει τότε οι εκτιμητές  $\widehat{\beta}_0$  και  $\widehat{\beta}_1$ , των ελαχίστων τετραγώνων δεν είναι BLUE (Best Linear Unbiased Estimator). Το θεώρημα Gauss - Markov μας λέει ότι αν ισχύουν οι συνθήκες παλινδρόμησης τότε οι εκτιμητές ελαχίστων τετραγώνων  $\hat{\alpha}$  και  $\hat{\beta}$ :

- Είναι αμερόληπτοι (unbiased) εκτιμητές των συντελεστών παλινδρόμησης του πληθυσμού, δηλαδή  $E(\hat{\alpha}) = \alpha$  και  $E(\hat{\beta}) = \beta$
- Είναι συνεπείς (consistent) καθώς με την επ' άπειρον αύξηση του δείγματος, οι εκτιμητές συγκλίνουν προς τους συντελεστές παλινδρόμησης του πληθυσμού
- Είναι αποτελεσματικοί (efficient) διότι έχουν τη μικρότερη διακύμανση άρα και το μικρότερο τυπικό σφάλμα (standard error) εκτίμησης μεταξύ όλων των αμερόληπτων εκτιμητών
- Το άθροισμα των καταλοίπων (ei) γύρω από την γραμμή παλινδρόμησης  $Y_i$  ισούται με το μηδέν

$$\sum e_i = \sum (Y_i - \bar{Y})$$

- Το κατάλοιπο ( $e_i$ ), μετρά την απόκλιση της εξαρτημένης μεταβλητής  $Y_i$  από την γραμμή παλινδρόμησης
- Ο διαταρακτικός όρος ( $\varepsilon$ ), μετρά την απόκλιση της εξαρτημένης μεταβλητής  $Y_i$  από την μέση τιμή της  $\bar{Y}$ .

Ένα σημαντικό ακόμη χαρακτηριστικό της ανάλυσης παλινδρόμησης είναι και ο συντελεστής παλινδρόμησης. Πιο συγκεκριμένα, μέρος της μεταβλητότητας που παρατηρείται στις τιμές της  $Y$  οφείλεται στις μεταβολές της  $X$  και μέρος στις επιδράσεις των τυχαίων παραγόντων ( $u$ ). Πόση μεταβλητότητα της  $Y$  εξηγείται από την  $X$  και πόση από τυχαίους παράγοντες;

Μέτρο του βαθμού προσαρμογής της ευθείας παλινδρόμησης στις παρατηρήσεις του δείγματος. Συμβολίζεται με  $R^2$ . Ο αριθμητικός μέσος μιας μεταβλητής είναι η καλύτερη πρόβλεψη όταν η μόνη διαθέσιμη πληροφορία είναι οι τιμές της ίδιας της μεταβλητής. Η  $X$  μπορεί να θεωρηθεί ότι ερμηνεύει την  $Y$  στον βαθμό που συμβάλλει στην πρόβλεψή της πέρα από τον μέσο.

Οι τύποι για τον υπολογισμό του συντελεστή προσδιορισμού είναι:

$$R^2 = \frac{RSS}{TSS} = 1 - \frac{ESS}{TSS}$$

ή

$$R^2 = \frac{\sum(\hat{Y}_i - \bar{Y})^2}{\sum(Y_i - \bar{Y})^2} = 1 - \frac{\sum(Y_i - \hat{Y}_i)^2}{\sum(Y_i - \bar{Y})^2}$$

Ο συντελεστής συσχέτισης είναι ένα ακόμη μέτρο που χαρακτηρίζει την ευθεία αλλά και την ανάλυση παλινδρόμησης και αντιπροσωπεύει τον βαθμό γραμμικής συσχέτισης ανάμεσα σε δύο μεταβλητές χωρίς να ενδιαφέρεται για την αιτιώδη σχέση που μπορεί να υπάρχει μεταξύ τους. Συμβολίζεται με το  $\rho$  και συνδέεται άμεσα με τον συντελεστή προσδιορισμού αφού είναι η τετραγωνική ρίζα αυτού:

$$\rho = \sqrt{R^2}$$

$$\rho = \frac{\text{cov}(X, Y)}{\sqrt{\text{var}(X)}\sqrt{\text{var}(Y)}}$$

## 1.5. ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ στο ΑΠΛΟ ΓΡΑΜΜΙΚΟ ΜΟΝΤΕΛΟ



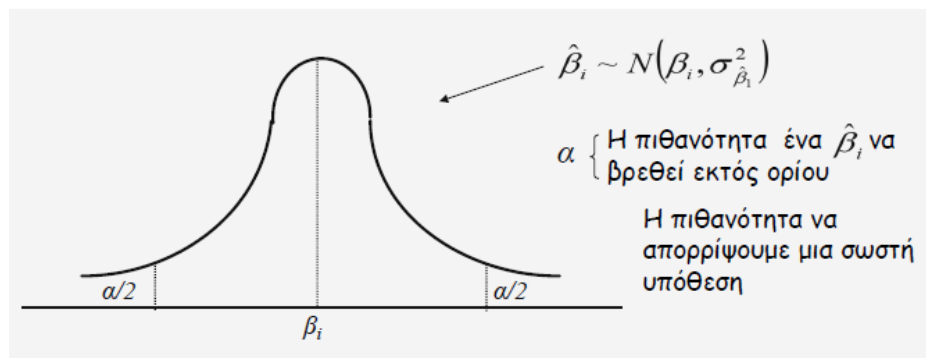
Πριν προσδιορίσουμε την έννοια των ελέγχων είναι χρήσιμο να αναφερθούμε στον προσδιορισμό των υποθέσεων, της μηδενικής και της εναλλακτικής. Πιο συγκεκριμένα:

- η μηδενική υπόθεση ( $H_0$ ) αναφέρεται σε μια τιμή ή σε ένα διάστημα τιμών ενός συντελεστή που συνήθως δεν είναι οι αναμενόμενες.
- η εναλλακτική υπόθεση ( $H_1$ ) αναφέρεται στην τιμή ή στο διάστημα τιμών του συντελεστή που ισχύουν αν δεν γίνει αποδεκτή η μηδενική υπόθεση.

Τα κριτήρια αποδοχής ή απόρριψης της μηδενικής υπόθεσης διαμορφώνονται ως εξής:

- **Περιοχή αποδοχής:** Το απαραίτητο εύρος τιμών της στατιστικής για να γίνει αποδεκτή η  $H_0$
- **Περιοχή απόρριψης:** Το απαραίτητο εύρος τιμών της στατιστικής για να απορριφθεί η  $H_0$

Πιο συγκεκριμένα τα παραπάνω αποτυπώνονται στο ακόλουθο διάγραμμα:



Σχήμα 6: Αποδοχή – Απόρριψη της Μηδενικής Υπόθεσης  $H_0$ <sup>5</sup>

Όσον αφορά την οικονομετρία, η αξιολόγηση του υποδείγματος:

5

<http://androulakis.bma.upatras.gr/mediawiki/index.php?title=%CE%A3%CE%91%CE%A3%CE%91%CE%9B%CE%9F%CE%A5%CE%95%CE%A5%CE%A6%CE%A1%CE%9F%CE%A3%CE%A5%CE%9D%CE%97&printable=yes>

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$$

Βασίζεται στην ακόλουθη μεθοδολογία:

- Οι συντελεστές  $\beta_i$  ελέγχονται με βάση την  $t$  – κατανομή
- Το σύνολο του υποδείγματος ελέγχεται με βάση την  $f$  – κατανομή
- Η σημαντικότητα της διακύμανσης του διαταρακτικού όρου ελέγχεται με βάση την  $X^2$  κατανομή.

Αν η πιθανότητα να συμβεί το ψευδές είναι  $\alpha$ , τότε το ενδεχόμενο «αληθές» έχει  $(1-\alpha)$  πιθανότητες να συμβεί. Το  $(1 - \alpha)$  ονομάζεται διάστημα εμπιστοσύνης (confidence interval) και είναι αυτό που προσδιορίζει την περιοχή εντός της οποίας είναι πιθανό να βρίσκονται οι τιμές των συντελεστών. Το  $\alpha$  καλείται επίπεδο σημαντικότητας (level significance) και είναι αυτό που ορίζει με ακρίβεια την περιοχή του διαστήματος εμπιστοσύνης.

Αρχικά πραγματοποιούμε ελέγχους για τη στατιστική σημαντικότητα των παραμέτρων του υποδείγματος, οι έλεγχο αυτοί διαμορφώνονται ως εξής:

#### **Δίπλευρος Έλεγχος:**

Στον δίπλευρο έλεγχο σκοπός μας είναι η διερεύνηση της παρακάτω υποθέσεως: Εάν υπάρχει στατιστικά σημαντική σχέση ανάμεσα στην εξαρτημένη και στην ανεξάρτητη μεταβλητή. Με άλλα λόγια εάν η κλίση της συνάρτησης παλινδρόμησης μας θα είναι θετική ή αρνητική με βάση την συναρτησιακή σχέση ανάμεσα στα  $X, Y$ . Ο έλεγχος διαμορφώνεται ως εξής:

$$H_0: \beta_1 = 0 \text{ (Μηδενική Υπόθεση)}$$

$$H_1: \beta_1 \neq 0 \text{ (Εναλλακτική Υπόθεση)}$$

#### **Δεξιός Μονόπλευρος Έλεγχος:**

Ας υποθέσουμε ότι με βάση την οικονομική θεωρία αναμένουμε ότι ο συντελεστής  $\beta_1$  είναι θετικός και θέλουμε να διαπιστώσουμε εμπειρικά αν ο  $\beta_1$  είναι, πρώτο, θετικός και, δεύτερο, στατιστικά σημαντικός.

Σε αυτή την περίπτωση, οι  $H_0$  και  $H_1$ , υποθέσεις εκφράζονται ως εξής:

$$H_0: \beta_1 = 0 \text{ (Μηδενική Υπόθεση)}$$

$$H_1: \beta_1 > 0 \text{ (Εναλλακτική Υπόθεση)}$$

#### **Αριστερός Μονόπλευρος Έλεγχος:**

Ας υποθέσουμε ότι με βάση την οικονομική θεωρία αναμένουμε ότι ο συντελεστής  $\beta_1$  είναι αρνητικός και θέλουμε να διαπιστώσουμε εμπειρικά αν ο  $\beta_1$  είναι, πρώτο, θετικός και, δεύτερο, στατιστικά σημαντικός.

Σε αυτή την περίπτωση, οι  $H_0$  και  $H_1$ , υποθέσεις εκφράζονται ως εξής:

$$H_0: \beta_1 = 0 \text{ (Μηδενική Υπόθεση)}$$

$$H_1: \beta_1 < 0 \text{ (Εναλλακτική Υπόθεση)}.$$

Τα πιθανά λάθη που μπορούν να προκύψουν κατά την διαδικασία του στατιστικού ελέγχου είναι:

- **Λάθος Τύπου I (type I error):** Αφορά την πιθανότητα κατά την οποία μετά τη διαδικασία του στατιστικού ελέγχου απορρίπτεται μια αληθής μηδενική υπόθεση
- **Λάθος Τύπου II (type II error):** Αφορά την πιθανότητα μετά την φάση του στατιστικού ελέγχου να γίνει αποδεκτή μια ψευδής μηδενική υπόθεση.

Στις υποενότητες που ακολουθούν περιγράφονται οι έλεγχοι υποθέσεων που γίνονται κάθε φορά προκειμένου να διαπιστωθεί αν οι παραπάνω υποθέσεις ισχύουν ή όχι. Πιο συγκεκριμένα:

### **1.5.1. ΕΤΕΡΟΣΚΕΛΑΣΤΙΚΟΤΗΤΑ**

Η ετεροσκεδαστικότητα οφείλεται σε διάφορες αιτίες. Οι πιο σημαντικές από αυτές είναι:

Η ετεροσκεδαστικότητα μπορεί να είναι μια φυσική ιδιότητα των μεταβλητών του υποδείγματος. Έτσι αν κατασκευάσουμε ένα οικονομετρικό υπόδειγμα με ανεξάρτητη μεταβλητή τις ώρες εξάσκησης στη δακτυλογράφηση και με εξαρτημένη μεταβλητή τον αριθμό των τυπογραφικών λαθών τότε αναμένουμε ότι όσο αυξάνουν οι ώρες μειώνεται όχι μόνο ο αριθμός των λαθών αλλά και η μεταβλητότητα αυτών. Επίσης αν κατασκευάσουμε ένα οικονομετρικό υπόδειγμα με ανεξάρτητη μεταβλητή τα ετήσια κέρδη των επιχειρήσεων και εξαρτημένη τα μερίσματα που δίνει η μετοχή των εταιριών και πάλι αναμένουμε ότι όσο αυξάνουν τα κέρδη θα αυξάνει τόσο το μέγεθος του μερίσματος όσο και η μεταβλητότητα αυτού. Στα παραπάνω παραδείγματα η δεσμευμένη διασπορά της εξαρτημένης μεταβλητής, και άρα και του διαταρακτικού όρου, δεν είναι σταθερή κατά μήκος των τιμών της ανεξάρτητης μεταβλητής, δηλαδή τα υποδείγματα παρουσιάζουν δεσμευμένη ετεροσκεδαστικότητα.

Η ετεροσκεδαστικότητα μπορεί να οφείλεται και σε ακραίες παρατηρήσεις (outliers) των μεταβλητών. Για παράδειγμα οποιοδήποτε οικονομετρικό υπόδειγμα με εξαρτημένη μεταβλητή τη μεταβλητότητα (volatility) των αποδόσεων του δείκτη S&P 500 συμπεριλάβει στο δείγμα παρατηρήσεις από το φθινόπωρο του 2008 θα παρουσιάζει δεσμευμένη ετεροσκεδαστικότητα. Τις ημέρες εκείνες η μεταβλητότητα έφθασε στο πρωτοφανές 80% όταν η μέση μεταβλητότητα της προηγούμενης δεκαετίας ήταν 20%.

Τέλος η ετεροσκεδαστικότητα μπορεί να οφείλεται και στο ότι το υπόδειγμα είναι λάθος εξειδικευμένο. Αυτό σημαίνει ότι μπορεί να απουσιάζει από αυτό μια σημαντική ανεξάρτητη μεταβλητή ή ότι η συναρτησιακή μορφή των μεταβλητών δεν είναι σωστή. Για παράδειγμα αν το σωστά εξειδικευμένο υπόδειγμα περιλαμβάνει ως ανεξάρτητη μεταβλητή τη  $x^2$  ενώ εμείς χρησιμοποιήσουμε τη  $x$  τότε θα παρουσιαστεί ετεροσκεδαστικότητα στα κατάλοιπα της παλινδρόμησης.

Έτσι λοιπόν όταν το υπόδειγμα παρουσιάζει δεσμευμένη ετεροσκεδαστικότητα οι εκτιμητές ελαχίστων τετραγώνων δεν είναι αποτελεσματικοί και επίσης δεν μπορούμε να διεξάγουμε ελέγχους υποθέσεων βασισμένοι σε αυτούς. Χρειαζόμαστε λοιπόν στη

περίπτωση αυτή να εισάγουμε κάποιους νέους εκτιμητές για το υπόδειγμα. Οι εκτιμητές αυτοί ονομάζονται εκτιμητές σταθμισμένων ελαχίστων τετραγώνων (weighted least squares) ή εκτιμητές γενικευμένων ελαχίστων τετραγώνων (generalized least squares). Η βασική ιδέα για τη κατασκευή αυτών των εκτιμητών είναι η μετατροπή του γραμμικού υποδείγματος σε ένα νέο το οποίο δεν παρουσιάζει ετεροσκεδαστικότητα και η εκτίμηση του νέου υποδείγματος με τη μέθοδο ελαχίστων τετραγώνων.

Συνοψίζοντας τα αποτελέσματα μέχρι τώρα για ένα υπόδειγμα που παρουσιάζει ετεροσκεδαστικότητα έχουμε δύο συνεπείς και ασυμπτωτικά κανονικούς εκτιμητές αυτού. Τους εκτιμητές ελαχίστων τετραγώνων και τους εκτιμητές εφικτών γενικευμένων ελαχίστων τετραγώνων. Γενικά ένας συνεπής και ασυμπτωτικά κανονικός εκτιμητής είναι ασυμπτωτικά πιο αποτελεσματικός από κάποιον άλλον επίσης συνεπή και ασυμπτωτικά κανονικό εκτιμητή αν η ασυμπτωτική διασπορά του δεύτερου είναι μεγαλύτερη ή ίση με αυτή του πρώτου. Μπορούμε να αποδείξουμε ότι οι εκτιμητές εφικτών γενικευμένων ελαχίστων τετραγώνων είναι ασυμπτωτικά πιο αποτελεσματικοί από τους εκτιμητές ελαχίστων τετραγώνων. Προσοχή όμως το συμπέρασμα αυτό βασίζεται στην υπόθεση ότι το μέγεθος του δείγματος είναι μεγάλο και ότι το υπόδειγμα της ετεροσκεδαστικότητας είναι σωστά εξειδικευμένο. Αν ένα από αυτά τα δύο δεν ικανοποιείται είναι πιθανό η ασυμπτωτική διασπορά των εκτιμητών εφικτών γενικευμένων ελαχίστων τετραγώνων να είναι μεγαλύτερη από αυτή των εκτιμητών ελαχίστων τετραγώνων.

Όλοι οι έλεγχοι ακολουθούν την ακόλουθη κοινή στρατηγική. Τα κατάλοιπα ελαχίστων τετραγώνων είναι συνεπείς εκτιμητές των τιμών του διαταρακτικού όρου ακόμα και όταν υπάρχει ετεροσκεδαστικότητα στο υπόδειγμα. Έτσι οι τιμές των καταλοίπων θα αντικατοπτρίζουν σε κάποιο βαθμό την πιθανή ετεροσκεδαστικότητα του υποδείγματος ή διαφορετικά, δεδομένης της δεσμευμένης ομοσκεδαστικότητας για το πληθυσμό πόσο πιθανό είναι να έχουμε ένα δείγμα με κατάλοιπα αυτά που έχουμε εκτιμήσει;

**Ο έλεγχος White** είναι ο γενικότερος έλεγχος δεσμευμένης ετεροσκεδαστικότητας που έχει παρουσιαστεί μέχρι στιγμής στη βιβλιογραφία. Ο έλεγχος White είναι πολύ γενικός μιας και δεν απαιτείται να κάνουμε κάποια

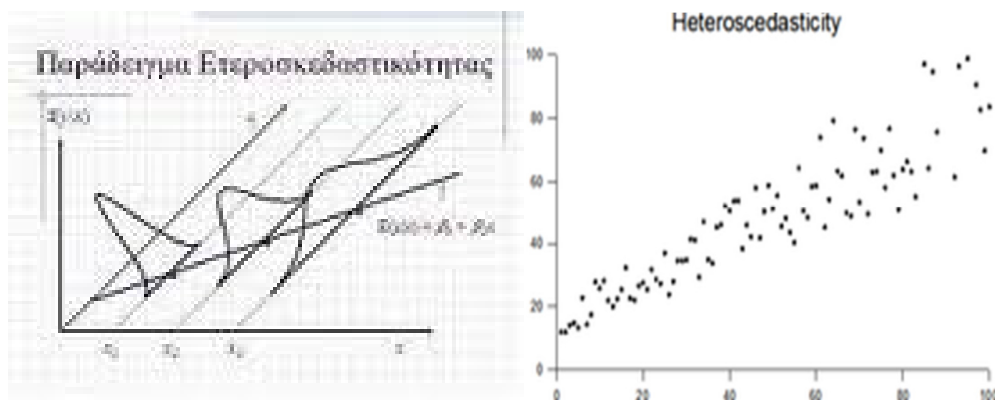
συγκεκριμένη υπόθεση για τη μορφή της ετεροσκεδαστικότητας. Το χαρακτηριστικό αυτό αποτελεί ταυτόχρονα πλεονέκτημα και μειονέκτημα του ελέγχου. Αν ο έλεγχος απορρίψει τη μηδενική υπόθεση αυτό μπορεί πολύ απλά να οφείλεται, όχι στο ότι υπάρχει ετεροσκεδαστικότητα, αλλά στο ότι το υπόδειγμα δεν είναι σωστά εξειδικευμένο (για παράδειγμα η παράλειψη της μεταβλητής  $x^2$  από τη παλινδρόμηση). Η ισχύς του ελέγχου White τείνει στη μονάδα όταν το  $N$  τείνει στο άπειρο, αλλά επειδή το διάλυσμα  $\psi$  μπορεί να έχει ένα μεγάλο αριθμό στοιχείων μπορεί να απαιτηθεί ένα αρκετά μεγάλο δείγμα για να επιτευχθεί η σύγκλιση αυτή.

Αυτό σημαίνει ότι όταν το υπόδειγμα έχει ένα μεγάλο αριθμό ανεξάρτητων μεταβλητών και το δείγμα είναι πεπερασμένο η ισχύς του ελέγχου μπορεί να αποκλίνει σημαντικά από τη μονάδα.

Στον έλεγχο **Goldfeld - Quandt** χωρίζουμε τις παρατηρήσεις σε δύο σύνολα με τέτοιο τρόπο έτσι ώστε υπό τη μηδενική υπόθεση της ομοσκεδαστικότητας η διασπορά θα είναι ίδια στα δύο σύνολα ενώ υπό την εναλλακτική θα είναι στατιστικά διαφορετική. Ο έλεγχος Goldfeld - Quandt απαιτεί το διαχωρισμό του δείγματος σε δύο υποσύνολα για τα οποία μπορούμε να υποθέσουμε ότι έχουν διαφορετική διασπορά. Αυτό όμως δεν είναι πάντα δυνατό σε όλα τα υποδείγματα.

Ο έλεγχος **Breusch-Pagan-Godfrey** βασίζεται στην παραμετροποίηση της διασποράς. Ουσιαστικά ο έλεγχος Breusch-Pagan-Godfrey είναι μια ειδική περίπτωση του ελέγχου White.

Γραφικά η ετεροσκεδαστικότητα απεικονίζεται ως εξής:



Σχήμα 7: Ετεροσκεδαστικότητα<sup>6</sup>

Σύμφωνα με τα παραπάνω, το πρόβλημα της ετεροσκεδαστικότητας, παρουσιάζεται συνήθως σε διαστρωματικά στοιχεία και το πρόβλημα της αυτοσυσχέτισης σε διαχρονικά στοιχεία, παρόλα αυτά υπάρχουν περιπτώσεις που η ετεροσκεδαστικότητα απαντάται και σε διαχρονικά στοιχεία.

Ερευνητές στην προσπάθειά τους να κατασκευάσουν υποδείγματα πρόβλεψης για χρηματοοικονομικά στοιχεία παρατήρησαν ότι σε διάφορες χρονικές περιόδους οι μεταβλητές παρουσιάζουν μεγάλη μεταβλητικότητα.

Αν προσπαθήσουμε να εκτιμήσουμε ένα τέτοιο υπόδειγμα πρόβλεψης θα καταλήξουμε ότι σε ορισμένες περιόδους τα σφάλματα προβλέψεων θα είναι μεγάλα (ασταθείς περίοδοι) και σε άλλες περιόδους μικρά (ήρεμοι περίοδοι), δηλαδή οι διακυμάνσεις των σφαλμάτων έτειναν να ομαδοποιούνται διαχρονικά κατά μέγεθος παρουσιάζοντας ένα είδος ετεροσκεδαστικότητας υπό συνθήκη.

Με τον τρόπο αυτό δημιουργήθηκαν ορισμένα υποδείγματα που λαμβάνουν υπόψιν τους τις διακυμάνσεις αυτές των διαταρακτικών όρων

Τα υποδείγματα αυτά είναι τα παρακάτω:

- Το υπόδειγμα ARCH
- Το υπόδειγμα GARCH

<sup>6</sup> <http://mathimataoikonometrias.blogspot.gr/2012/02/blog-post.html>

- Το υπόδειγμα GARCH – M.

### 1.5.2. ΑΥΤΟΣΥΣΧΕΤΙΣΗ

Οι βασικότερες αιτίες που προκαλούν αυτοσυσχέτιση είναι οι ακόλουθες:

Οι περισσότερες οικονομικές χρονολογικές σειρές, όπως για παράδειγμα το ΑΕΠ, η κατανάλωση, η απασχόληση κ.α, παρουσιάζουν αδράνεια. Αυτό σημαίνει ότι κινούνται με βάση οικονομικούς κύκλους. Ξεκινώντας από τα χαμηλά μιας περιόδου ύφεσης οι τιμές αυτών των μεταβλητών μεταβάλλονται θετικά, δηλαδή η τιμή της επόμενης περιόδου είναι μεγαλύτερη από τη τιμή της προηγούμενης. Το φαινόμενο αυτό συνεχίζεται μέχρις ότου ξαναεμφανιστεί ύφεση και η μεταβολή επιβραδυνθεί ή και αναστραφεί.

Το φαινόμενο της αυτοσυσχέτισης μπορεί επίσης να οφείλεται στο ότι απουσιάζουν κάποιες ανεξάρτητες μεταβλητές από το υπόδειγμα οι οποίες επηρεάζουν την εξαρτημένη μεταβλητή ή ότι το υπόδειγμα δεν έχει τη σωστή συναρτησιακή μορφή.

Η τρέχουσα τιμή της εξαρτημένης μεταβλητής μπορεί να εξαρτάται, πέραν όλων των άλλων, και από τις παρελθούσες τιμές αυτής. Αν λοιπόν οι παρελθούσες τιμές απουσιάζουν από τις ανεξάρτητες μεταβλητές το υπόδειγμα θα παρουσιάσει αυτοσυσχέτιση.

Οι έλεγχοι αυτοσυσχέτισης βασίζονται στην ιδέα ότι αν ο διαταρακτικός όρος παρουσιάζει αυτοσυσχέτιση τότε αυτή θα εμφανιστεί και στα κατάλοιπα ελαχίστων τετραγώνων.

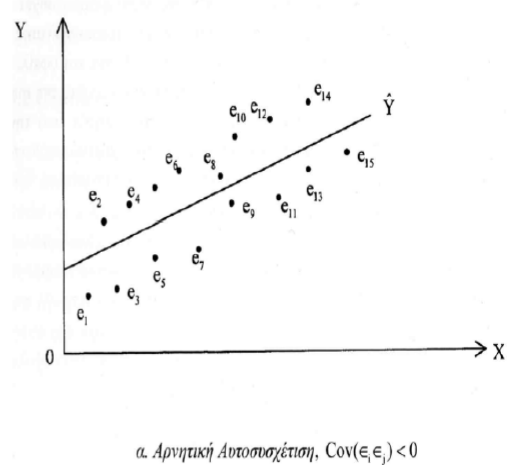
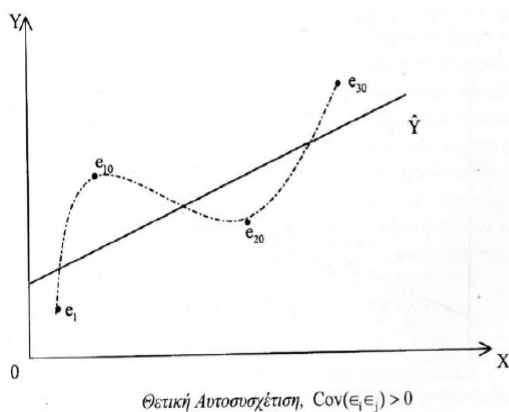
Ο έλεγχος **Durbin-Watson** ήταν ο πρώτος που παρουσιάστηκε για να εξετάσει πιθανή ύπαρξη αυτοσυσχέτισης στο διαταρακτικό όρο. Σε αντίθεση με όλους τους άλλους ελέγχους η κατανομή δειγματοληψίας της στατιστικής Durbin-Watson εξαρτάται από τις ανεξάρτητες μεταβλητές  $X$ . Αυτό συνεπάγεται ότι και οι κριτικές τιμές του ελέγχου εξαρτώνται από τις ανεξάρτητες μεταβλητές. Όμως οι Durbin και Watson έδειξαν ότι η οι κριτικές τιμές φράσσονται αριστερά και δεξιά από δύο μεγέθη τα οποία



εξαρτώνται μόνο από το μέγεθος του δείγματος, τον αριθμό των ανεξάρτητων μεταβλητών και φυσικά το επίπεδο σημαντικότητας  $\alpha$ .

Ο έλεγχος **Durbin - Watson** απαιτεί οι διαταρακτικός όρος να ακολουθεί κανονική κατανομή και μπορεί να ελέγξει μόνο για αυτοσυσχέτιση 1ης τάξης. Αν θέλουμε να κατασκευάσουμε ελέγχους οι οποίοι θα γενικεύουν αυτά τα χαρακτηριστικά θα πρέπει να στραφούμε στην ασυμπτωτική θεωρία.

Οι έλεγχοι **Box - Pierce** και **Ljung - Box** υποθέτουν ότι  $E(\epsilon_i / X) = 0$ , το οποίο σημαίνει για παράδειγμα ότι ο πίνακας  $X$  δεν περιλαμβάνει παρελθούσες παρατηρήσεις της εξαρτημένης μεταβλητής. Αν όμως η τελευταία υπόθεση δεν ικανοποιείται από το γραμμικό υπόδειγμα τότε οι στατιστικές μπορεί να μην συγκλίνουν ασυμπτωτικά στην  $X^2$  κατανομή. Ο έλεγχος **Breusch - Godfrey** έρχεται να καλύψει το κενό αυτό των παραπάνω ελέγχων.



Σχήμα 8: Αυτοσυσχέτιση

Συμπερασματικά οι λόγοι που προκαλούν την αυτοσυσχέτιση:

- Η εσφαλμένη αλγεβρική εξειδίκευση του υποδείγματος
- Η παράλειψη μιας ή περισσότερων ερμηνευτικών μεταβλητών
- Η ύπαρξη συστηματικού σφάλματος μέτρησης στις μεταβλητές

- Η εκτίμηση μέρους των παρατηρήσεων με παρεμβολή
- Η κατανομή της επίδρασης ορισμένων τυχαίων γεγονότων σε περισσότερες από μια χρονικές περιόδους.

### 1.5.3. ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ

Ο όρος “πολυσυγγραμμικότητα” εισήχθη για πρώτη φορά από τον Ragnar Frisch (1934) και εκφράζει την ύπαρξη μίας ή περισσότερων, ταυτοχρόνως, γραμμικών σχέσεων μεταξύ των ανεξαρτήτων μεταβλητών που συναπαρτίζουν ένα οικονομετρικό υπόδειγμα. Η ταυτόχρονη χρησιμοποίηση, σε ένα πολλαπλό γραμμικό υπόδειγμα, πολλών ανεξαρτήτων μεταβλητών που συσχετίζονται γραμμικά μεταξύ τους δεν οδηγεί σε βελτίωση της πληροφορίας που εμπερικλείεται στο υπόδειγμα. Ως άμεση συνέπεια του γεγονότος αυτού είναι το να μην επιτυγχάνεται ο βέλτιστος προσδιορισμός των τιμών της εξαρτημένης μεταβλητής.

Η κατάσταση η οποία δημιουργείται όταν υπάρχουν ισχυρές συσχετίσεις μεταξύ των ανεξάρτητων μεταβλητών στην πολλαπλή παλινδρόμηση ονομάζεται πολυσυγγραμμικότητα (multicollinearity).

Στις περιπτώσεις που το πρόβλημα αυτό υφίσταται θα πρέπει κανείς να είναι ιδιαίτερα προσεκτικός στην ερμηνεία όλων των εκτιμητριών που προκύπτουν από το μοντέλο αυτό.

Υπάρχουν μια σειρά από προειδοποιητικές ενδείξεις που αν ο ερευνητής τις προσέξει είναι δυνατόν να αντιληφθεί ότι υπάρχει πολυσυγγραμμικότητα. Η πιο σημαντική από αυτές είναι ο πίνακας των συντελεστών συσχέτισης (Correlation Matrix) των ανεξάρτητων μεταβλητών. Αν στον πίνακα αυτόν υπάρχουν μεγάλες θετικές ή αρνητικές τιμές θα έχουμε μια ένδειξη ότι οιαντίστοιχες ανεξάρτητες μεταβλητές που χρησιμοποιούνται στο μοντέλο έχουν μεταξύ τους ισχυρό βαθμό συσχέτισης. Το στατιστικό συμπέρασμα που προκύπτει στις περιπτώσεις αυτές είναι ότι κάποιες από τις

μεταβλητές συνεισφέρουν ελάχιστα ή καθόλου στην πρόβλεψη της εξαρτημένης μεταβλητής οπότε και θα πρέπει να απομακρυνθούν από το μοντέλο.

Εαν, παρ' όλα αυτά, ο ερευνητής είναι βέβαιος ότι ο καθορισμός των ανεξάρτητων μεταβλητών έγινε σωστά θα πρέπει να εξετάσει δύο άλλες ενδείξεις για το κατά πόσον υπάρχει πολυσυγγραμμικότητα.

- Εάν τα πρόσημα ορισμένων συντελεστών στην παλινδρόμηση είναι αντίθετα από αυτά που θα περίμενε κανείς λόγω της φύσης του προβλήματος και
- Εάν σημαντικοί συντελεστές της παλινδρόμησης εμφανίζονται να έχουν μεγάλες τιμές στις τυπικές αποκλίσεις τους.

Οποιαδήποτε από τις δύο αυτές ενδείξεις θα πρέπει να προβληματίσει τον ερευνητή και να τον οδηγήσει σε μια σοβαρή έρευνα για το κατά πόσον υφίσταται πολυσυγγραμμικότητα.

Ο καθορισμός εκείνων των γραμμικών συνδυασμών των παραμέτρων  $\beta$  που μπορούν να εκτιμηθούν με ακρίβεια είναι εξαιρετικά δύσκολο στην περίπτωση που υφίσταται πολυσυγγραμμικότητα. Παρότι, εν γένει, δεν είναι δυνατόν να εξαληφθεί τελείως το πρόβλημα αυτό υπάρχει μια διαδικασία με την οποία ο ερευνητής μπορεί να εργασθεί με ένα μοντέλο που προκύπτει από το αρχικό με μετασχηματισμό των αρχικών μεταβλητών σε ένα σύνολο άλλων μεταβλητών που είναι ασυσχέτιστες μεταξύ τους. Η μεθοδολογία αυτή ονομάζεται ανάλυση κυρίων συνιστωσών (principal component analysis). Η τεχνική αυτή είναι μια παρά πολύ ισχυρή τεχνική τόσο στον εντοπισμό της πολυσυγγραμμικότητας όσο και ως μεθοδολογία που οδηγεί στον καθορισμό εκείνων των γραμμικών συνδυασμών των συντελεστών παλινδρόμησης που μπορούν να εκτιμηθούν με ακρίβεια.

Μια άλλη προσέγγιση είναι να υπολογισθούν οι  $k$  συντελεστές προσδιορισμού των παλινδρομήσεων κάθε μιας από τις ανεξάρτητες μεταβλητές στις υπόλοιπες  $k-1$  ανεξάρτητες μεταβλητές. Εκείνες οι μεταβλητές που εμφανίζουν υψηλό συντελεστή προσδιορισμού θα πρέπει να θεωρηθεί ότι είναι συγγραμμικές με τουλάχιστον μια από τις υπόλοιπες μεταβλητές. Στη συνέχεια θα πρέπει να υπολογισθεί η ομάδα εκείνων των μεταβλητών που έχουν υψηλή πολυσυγγραμμικότητα και μια ή περισσότερες από αυτές

τις μεταβλητές μέσα στη συγκεκριμένη ομάδα θα πρέπει να απομακρυνθούν πριν προχωρήσει κανείς σε ανάλυση παλινδρόμησης για την αρχική εξαρτημένη μεταβλητή.

Συνοψίζοντας να αναφέρουμε πως το πρόβλημα της πολυσυγγραμμικότητας επηρεάζει το εύρος των τιμών του διαστήματος εμπιστοσύνης των παραμέτρων ενός υποδείγματος καθώς και την αξιοπιστία των στατιστικών ελέγχων που διενεργούνται επί των εν λόγω παραμέτρων. Επηρεάζει, επίσης, την ακρίβεια και τη σταθερότητα των λοιπών εκτιμήσεων που λαμβάνουν χώρα επί του θεωρουμένου υποδείγματος. Επιπροσθέτως, δημιουργεί προβλήματα στην ερμηνεία των προκύπτοντων αποτελεσμάτων καθώς επίσης και στον καθορισμό του υποδείγματος.

#### **1.5.4. ΣΦΑΛΜΑ ΕΞΕΙΔΙΚΕΥΣΗΣ**

Συνήθως ο όρος σφάλμα εξειδίκευσης (specification error) αναφέρεται στα σφάλματα που δημιουργούνται από λαθεμένη διατύπωση της εξίσωσης παλινδρόμησης (παραλείπεται από το υπόδειγμα μια σημαντική ερμηνευτική μεταβλητή) ή στη χρησιμοποίηση λαθεμένης μορφής συνάρτησης (π.χ. γραμμική αντί εκθετική).

Ο γενικός έλεγχος διερεύνησης των σφαλμάτων εξειδίκευσης ενός υποδείγματος είναι ο έλεγχος RESET (Regression Specification Error Test) που προτάθηκε από τον Ramsey (1969). Ο έλεγχος αυτός χρησιμοποιεί τα τετράγωνα των εκτιμημένων τιμών της εξαρτημένης μεταβλητής και ακολουθεί τα παρακάτω βήματα:

##### **ΒΗΜΑ 1**

Γράφω τις δύο υποθέσεις για την κανονικότητα των καταλοίπων:

*H<sub>0</sub>: Το υπόδειγμα είναι σωστά εξειδικευμένο*

*H<sub>a</sub>: Το υπόδειγμα δεν είναι σωστά εξειδικευμένο*

Ο έλεγχος για την εξειδίκευση του υποδείγματος γίνεται με την F κατανομή (έλεγχος του Wald) καθώς και με  $X^2$  κατανομή (έλεγχος του λόγου πιθανοφανειών).

##### **ΒΗΜΑ 2**

Σχηματίζοντας την F κατανομή βρίσκω το κρίσιμο σημείο για επίπεδο σημαντικότητας 5% και βαθμούς ελευθερίας  $v_1 = h$  και  $v_2 = n - (k + 1 + h)$  ή την  $X^2$  κατανομή και βρίσκω το κρίσιμο σημείο για επίπεδο σημαντικότητας 5% και βαθμούς ελευθερίας  $v = h$

### **ΒΗΜΑ 3**

Εκτιμούμε το υπόδειγμα με τη μέθοδο των ελαχίστων τετραγώνων και σώζουμε τις εκτιμημένες τιμές καθώς και τον συντελεστή προσδιορισμού, όπως και την τιμή της πιθανοφάνειας  $\lambda_1$

### **ΒΗΜΑ 4**

Εισάγουμε στο υπόδειγμα παλινδρόμησης  $h$  επιπλέον ερμηνευτικές μεταβλητές οι οποίες αποτελούν δυνάμεις της εκτιμημένης μεταβλητής, δηλαδή εισάγουμε τις μεταβλητές  $, \dots$  και εκτιμούμε το νέο υπόδειγμα και σημειώνουμε το νέο συντελεστή προσδιορισμού, καθώς και τη νέα τιμή της πιθανοφάνειας  $\lambda_2$

### **ΒΗΜΑ 5**

Με τον έλεγχο του Wald ελέγχουμε αν οι επιπλέον ερμηνευτικές μεταβλητές είναι σημαντικές. Για τον έλεγχο αυτό χρησιμοποιούμε το στατιστικό F και την παρακάτω ποσότητα:

$$F = \frac{\frac{(R_2^2 - R_1^2)}{h}}{\frac{(1 - R_2^2)}{[n - (k + 1 + h)]}}$$

Ενώ για τον έλεγχο του λόγου πιθανοφάνειών, υπολογίζουμε το στατιστικό LR από την παρακάτω ποσότητα:  $LR = -2(\lambda_1 - \lambda_2)$

### **ΒΗΜΑ 6**

Αν ισχύει:

$$F > F_{\text{pin}}[h, n-(k+1+h)]$$

Τότε απορρίπτω την  $H_0$ .

Ομοίως, απορρίπτω την  $H_0$  και αν:

$$LR > X^2_{\text{pin}}(h)$$

### 1.5.5. ΟΛΟΚΛΡΩΣΗ - ΣΥΝΟΛΟΚΛΗΡΩΣΗ

Η παλινδρόμηση μη στάσιμων χρονολογικών σειρών μπορεί να οδηγήσει σε στατιστικά αξιόπιστα συμπεράσματα αν οι σειρές είναι συνολοκληρωμένες.

Ολοκλήρωση: η μετατροπή μίας μη-στάσιμης σειράς σε στάσιμη επιτυγχάνεται αν εκφράσουμε την σειρά σε διαφορές. Αν η σειρά πρέπει να εκφραστεί  $d$  φορές σε διαφορές για να γίνει στάσιμη λέμε ότι είναι ολοκληρωμένη σε  $d$  βαθμό και συμβολίζεται:

$$Y_t \rightarrow I(d)$$

Μία στάσιμη σειρά είναι  $I(0)$ .

Αν σε ένα οικονομετρικό υπόδειγμα υπάρχουν δύο χρονολογικές σειρές ολοκληρωμένες με βαθμό ολοκλήρωσης 1, δηλαδή:

$$Y_t \rightarrow I(1) \text{ και } X_t \rightarrow I(1)$$

Τότε οι χρονολογικές σειρές είναι συνολοκληρωμένες αν τα κατάλοιπα από την παλινδρόμηση της μίας στην άλλη είναι στάσιμα δηλαδή αν:

$$e_t \rightarrow I(0)$$

Ο έλεγχος για συνολοκλήρωση βασίζεται ακριβώς στην εξέταση της ύπαρξης μοναδιαίας ρίζας στα κατάλοιπα από την παλινδρόμηση των δύο σειρών (Engle-Granger test).

## 1.5. ΠΡΟΒΛΕΨΕΙΣ

Οι προβλέψεις αποτελούν αναπόσπαστο κομμάτι της οικονομετρικής ανάλυσης και αφορούν στον τρόπο με τον οποίο με τη χρήση του κατάλληλου υποδείγματος θα καταφέρει ο ερευνητής/ μελετητής να προβλέψει την τιμή του υποδείγματος στο μέλλον.

Η πρόβλεψη των μελλοντικών τιμών μιας χρονολογικής σειράς βασίζεται στα αποτελέσματα του εκτιμώμενου υποδείγματος. Τα αποτελέσματα αυτά έχουν να κάνουν βασικά με τις εκτιμήσεις για τα  $\alpha$ ,  $\beta_1, \dots, \beta_p$ ,  $\theta_1, \dots, \theta_q$ . Έτσι, μια πρόβλεψη που κάνουμε για την περίοδο  $T + 1$  εκφράζεται ως:

$$\hat{y}_{T+1} = E(y_{T+1} | y_1, \dots, y_T)$$

Επομένως, η αναμενόμενη τιμή την χρονολογικής μας σειράς για την περίοδο  $T + 1$  στηρίζεται στην πληροφόρηση μέχρι την περίοδο  $T$ .

Για την πραγματοποίηση προβλέψεων υπάρχουν και κάποιοι εκτιμητές τους οποίους θα πρέπει να λαμβάνουμε υπόψη μας, οι σημαντικότεροι εκ των οποίων είναι:

**Μέσο Σφάλμα Τετραγώνου (Mean Square Error):**

$$MSE = \frac{1}{N} \sum_{t=1}^N \hat{\epsilon}_t^2$$

**Τετραγωνική ρίζα του Mean Square Error:**

$$RMSE = \sqrt{\frac{1}{N} \sum_{t=1}^N \hat{\epsilon}_t^2}$$

**Μέσο Απόλυτο Σφάλμα (Mean Absolute Error):**

$$MSE = \frac{1}{N} \sum_{t=1}^N |\hat{\epsilon}_t|$$



## **ΚΕΦΑΛΑΙΟ 2 «ΑΝΑΛΥΣΗ ΧΡΟΝΟΣΕΙΡΩΝ – ΑΝΑΛΥΣΗ ΔΙΑΣΤΡΩΜΑΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ»**

Στο παρόν κεφάλαιο θα αναλυθούν εκτενώς οι χρονοσειρές αλλά και η ανάλυση διαστρωματικών στοιχείων.

Πιο συγκεκριμένα, η ανάλυση χρονοσειρών (time series analysis) είναι το πεδίο έρευνας που μελετά συστήματα, διαδικασίες, σήματα και πρότυπα που εξελίσσονται χρονικά. Η ανάλυση χρονοσειρών έχει δυο βασικούς στόχους:

- να μελετήσει και να αναγνωρίσει τη φύση ενός φαινομένου που αναπαρίσταται από μια ακολουθία παρατηρήσεων, και
- να προβλέψει τη μελλοντική εξέλιξη του φαινομένου, δηλαδή τις μελλοντικές τιμές της ακολουθίας παρατηρήσεων

Από την άλλη, στο δεύτερο μισό του κεφαλαίου θα αναλυθούν τα διαστρωματικά στοιχεία και η ανάλυση αυτών.

### **2.1. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ – ΧΡΗΣΙΜΟΤΗΤΑ - ΧΡΟΝΟΣΕΙΡΩΝ**

Με τον όρο χρονολογικές σειρές εννοούμε μια σειρά από παρατηρήσεις που λαμβάνονται σε ορισμένες χρονικές στιγμές ή περιόδους που ισαπέχουν ή μη μεταξύ τους. Έτσι, η τρέχουσα τιμή μιας μεταβλητής  $Y$  εκφράζεται ως συνάρτηση των προηγούμενων τιμών της, δηλαδή των τιμών της με χρονική υστέρηση.

Αρχικά θα μπορούσαμε να σημειώσουμε πως στόχος της στοχαστικής ανάλυσης των χρονολογικών σειρών είναι η μελέτη της στοχαστικής διαδικασίας διαμόρφωσης των δεδομένων (data generating process) και η διενέργεια προβλέψεων. Στην πράξη οι χρονολογικές σειρές στοχεύουν στην ενσωμάτωση της μεταβλητικότητας μιας χρονολογικής σειράς μέσω ενός υποδείγματος χρονολογικών σειρών. Η ενσωμάτωση πραγματοποιείται με τη λογική του μαύρου κουτιού (Black Box). Δηλαδή, επιλέγουμε το

υπόδειγμα που επιτυγχάνει τη καλύτερη δυνατή ενσωμάτωση με τις λιγότερες μεταβλητές. Αφου το επιτύχουμε αυτό. Πραγματοποιούμε μελλοντικές προβλέψεις μέσω του υποδείγματος το οποίο έχουμε επιλέξει.

Με άλλα λόγια, στόχος των χρονολογικών σειρών είναι, 'ενώνοντας' όλα τα γεγονότα που έχουν συλλεχτεί στον χρόνο, να δημιουργηθεί ένα χρονοδιάγραμμα, η μελέτη του οποίου θα δώσει μια γενική εικόνα της διαχρονικής εξέλιξης των φαινομένων ή χαρακτηριστικών. Με άλλα λόγια η ανάλυση χρονολογικών σειρών χρησιμοποιείται για να καθοριστούν μοντέλα που μετατρέπουν πληροφορίες από κανονικά χρονικά διαστήματα σε στατιστικά μέτρα. Οι κυριότερες μέθοδοι ανάλυσης χρονοσειρών είναι η μέθοδος της αυτοσυσχέτισης (auto correlation) όπου η χρονοσειρά αναπαριστάται με ένα δυναμικό μοντέλο (μοντέλο ARIMA, διαδικασία Box&Jenkins), όπου οι παρατηρήσεις θεωρούνται ως συναρτήσεις του παρελθόντος τους (και πιθανόν του παρελθόντος και άλλων μετρούμενων ή παρατηρούμενων μεταβλητών) και η μέθοδος της φασματικής ανάλυσης (spectral analysis) όπου αναπαριστά την χρονοσειρά με ένα κινητικό μοντέλο που οι παρατηρήσεις θεωρούνται συναρτήσεις του χρόνου.

Οι σκοποί της ανάλυσης χρονολογικών σειρών είναι η περιγραφή μιας χρονολογικής σειράς, η εξήγησή της, η πρόγνωση, ο σχεδιασμός, ο έλεγχος και τέλος η κατανόηση της διαδικασίας. Η περιγραφή μιας χρονολογικής σειράς μπορεί να γίνει με μια απλή παρατήρηση της γραφικής της παράστασης 'με το μάτι', όπου ανακαλύπτονται περιοδικότητες, τάσεις και άλλα.

Επίσης, υπάρχουν περισσότερο σύνθετες και προχωρημένες περιγραφές που γίνονται με την βοήθεια των στοχαστικών μοντέλων. Η εξήγηση με την σειρά της, είναι η συσχέτιση δυο ή περισσότερων χρονολογικών σειρών μεταξύ τους και η δυνατότητα η μια να εξηγεί την άλλη. Επίσης, σημαντικός σκοπός στην ανάλυση των χρονολογικών σειρών είναι ο έλεγχος. Παραδείγματα στην σπουδαιότητα του ελέγχου μπορούν να παρουσιαστούν στις διαδικασίες παραγωγής και ποιότητας ενός προϊόντος.

Τέλος, βασικός σκοπός της ανάλυσης χρονολογικών σειρών είναι η κατανόηση της διαδικασίας. Την κατανόηση παρέχει η στατιστική δίνοντας φορμαλιστικές περιγραφές των χρονοσειρών καθαυτών. Έτσι, μπορεί να παραχθεί μια χρονοσειρά από ένα άγνωστο

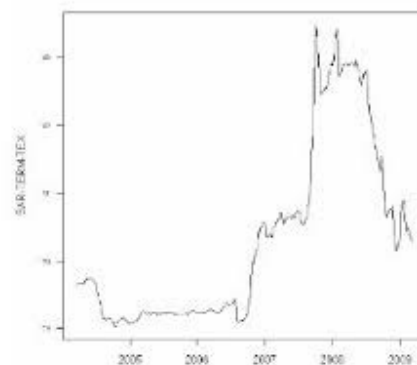
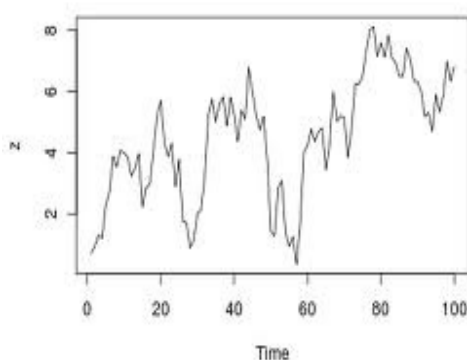
σύστημα και συνεπώς να κατανοηθεί καλύτερα το ίδιο το σύστημα που έχει παράξει την χρονοσειρά αυτή.

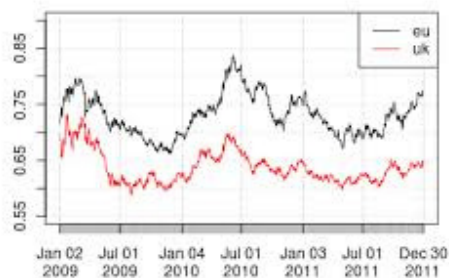
Παραδείγματα χρονολογικών σειρών μπορούν να παρουσιαστούν και στην Οικονομία. Τέτοια μπορεί να είναι η εξέλιξη της τιμής του πετρελαίου, η αξία μιας μετοχής, ο δείκτης συναλλαγών στο χρηματιστήριο και άλλα. Κύριο ενδιαφέρον για την οικονομία είναι η πρόβλεψη.

Στο εμπόριο (marketing) επίσης είναι οι πωλήσεις ανά μήνα. Βασικός σκοπός στο εμπόριο εκτός της πρόβλεψης είναι ο σχεδιασμός, η σύγκριση και η ερμηνεία. Επίσης, παραδείγματα χρονολογικών σειρών έχουμε από την δημογραφία όπου μελετώνται οι εξελίξεις πληθυσμών. Από την ιατρική αναλύονται τα ηλεκτροεγκεφαλογράφημα (ΗΕΓ, EEG) και ηλεκτροκαρδιογραφήματα (ΗΚΓ, ECG).

Ακόμα, χρονολογικές σειρές συναντάμε στην φυσική, στην σεισμολογία, στην μετεωρολογία, στην αστροφυσική και αλλού, με αντικειμενικό στόχο την κατανόηση και την πρόβλεψη. Στην αστροφυσική με την ανάλυση χρονοσειρών μπορεί να περιγραφεί το μοντέλο και να γίνει σύγκριση με τα θεωρητικά μοντέλα που ήδη υπάρχουν. Έτσι, με την κατανόηση των διαδικασιών που παράγουν τις χρονοσειρές καθοδηγούμαστε στην δημιουργία νέων μοντέλων. Τα νέα αυτά μοντέλα χρησιμοποιούνται μετά για προβλέψεις.

Μερικές εικόνες για τη μορφή των χρονολογικών σειρών είναι οι ακόλουθες:





Σχήμα 9: Χρονολογικές Σειρές<sup>7</sup>

Στην πράξη η ανάλυση χρονολογικών σειρών πραγματοποιεί τα εξής:

- Ενσωμάτωση της μεταβλητικότητας μιας χρονολογικής σειράς μέσω ενός υποδείγματος χρονολογικών σειρών
- Η ενσωμάτωση πραγματοποιείται με τη λογική του μαύρου κουτιού (Black Box)
- Επιλέγουμε το υπόδειγμα που επιτυγχάνει τη καλύτερη δυνατή ενσωμάτωση με τις λιγότερες μεταβλητές
- Αφού το επιτύχουμε αυτό. Πραγματοποιούμε μελλοντικές προβλέψεις μέσω του υποδείγματος το οποίο έχουμε επιλέξει.

## 2.2. ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ & ΜΕΤΡΑ ΧΡΟΝΟΛΟΓΙΚΩΝ ΣΕΙΡΩΝ

Οι χρονολογικές σειρές διακρίνονται σε **συνεχείς** χρονολογικές σειρές και σε **διακριτές**.

**Συνεχείς (continuous) χρονολογικές σειρές** είναι αυτές όπου η τιμή του φαινομένου παρατηρείται συνεχώς. Παράδειγμα συνεχών χρονολογικών σειρών είναι η

<sup>7</sup>

<https://www.google.gr/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0CAYQjB0&url=http%3A%2F%2Fstavrakoudis.econ.uoi.gr%2Fstavrakoudis%2F%3Fiid%3D204&ei=bueEVPiABMn8UtzIgNgP&bvm=bv.80642063.d.d24&psig=AFQjCNFJWl3sj-5kkBORjpK45heQu6SBRQ&ust=1418082534910173>

συνεχόμενη καταγραφή της θερμοκρασίας του αέρα ή η συνεχής παρακολούθηση των σεισμών.

**Διακριτές (discrete) χρονολογικές σειρές** είναι αυτές όπου η τιμή του φαινομένου καταγράφεται σε ορισμένα χρονικά διαστήματα. Παράδειγμα διακριτών χρονολογικών σειρών είναι η τιμή μιας μετοχής ανά ημέρα ή ο αριθμός των ηλιακών κηλίδων ανά έτος όπου υπάρχουν τιμές σε συγκεκριμένα χρονικά διαστήματα.

Οι διακριτές χρονολογικές σειρές είναι αυτές που μπορούν να 'κατανοηθούν' από έναν Η/Υ. Συνεπώς, αντικειμενικός στόχος είναι οι συνεχείς χρονοσειρές να μετατραπούν σε διακριτές.

Άλλες βασικές έννοιες των χρονολογικών σειρών είναι η γραφική παράσταση, ο μέσος όρος και η διασπορά αυτής, ο κινητό μέσος όρος, η τάση, η στασιμότητα και άλλα.

Με την έννοια **γραφική παράσταση** μιας χρονολογικής σειράς εννοούμε την καμπύλη που παράγεται σε ένα σύστημα ορθογωνίων αξόνων όπου ο οριζόντιος άξονας είναι ο χρόνος και ο κατακόρυφος οι μετρούμενες τιμές στο αντίστοιχο χρονικό διάστημα.

**Μέσος όρος** μιας χρονολογικής σειράς είναι η μέση τιμή όλων των τιμών της.

**Διασπορά** μιας χρονολογικής σειράς είναι ο μέσος όρος των τετραγώνων των αποκλίσεων από την μέση τιμή.

**Τυπική απόκλιση** μιας χρονολογικής σειράς είναι η τετραγωνική ρίζα της διασποράς. Η τυπική απόκλιση είναι περισσότερο διαισθητική έννοια από ότι η διασπορά.

Ορισμένες από τις βασικότερες συνιστώσες μιας χρονολογικής σειράς είναι οι ακόλουθες:

**Μακροχρόνια Τάση (trend) -  $T_t$**

- γραμμική
- μη - γραμμική

**Κυκλική περιοδικότητα (cyclical variation) -  $C_t$**

Αφορά αυξήσεις και μειώσεις των τιμών της σειράς (αποκλίσεις δηλαδή από την μακροχρόνια τάση της) ανά διαστήματα μεγαλύτερα του έτους.

#### **Εποχιακή περιοδικότητα (seasonal variation) - $S_t$**

Η συμπεριφορά των τιμών της σειράς εναλλάσσεται σε διαστήματα μικρότερα του έτους (εβδομάδα, μήνας, τρίμηνο)

#### **Τυχαία περιοδικότητα (residual variation) - $R_t$**

$$Y_t = T_t + C_t + S_t + R_t$$

Οι βασικές κατηγορίες υποδειγμάτων χρονολογικών σειρών είναι οι ακόλουθες, (ορισμένες από αυτές θα τις μελετήσουμε εκτενέστερα σε ακόλουθες ενότητες):

- Υποδείγματα Λευκού Θορύβου (white noise models)
- Υποδείγματα Τυχαίας Διαδρομής (random walk models)
- Αυτοπαλίνδρομα Υποδείγματα (autoregressive models-AR)
- Υποδείγματα Κινητών Μέσων (moving average models)
- Αυτοπαλίνδρομα Υποδείγματα Κινητών Μέσων (ARMA models)
- Υποδείγματα Box-Jenkins
- Vector Autoregression (VAR) Υποδείγματα

### **2.3. ΣΤΑΣΙΜΕΣ – ΜΗ ΣΤΑΣΙΜΕΣ ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ**

Μια στοχαστική διαδικασία είναι αυστηρώς ή πλήρως στάσιμη (strictly – strongly – completely stationary) όταν οι ιδιότητες της δεν επηρεάζονται από μια αλλαγή στην αρχή μετρήσεως του χρόνου. Αυτό σημαίνει ότι η συνδυασμένη συνάρτηση πιθανότητας με αρχή το χρονικό σημείο  $t$ , είναι ακριβώς η ίδια με τη συνδυασμένη συνάρτηση πιθανότητας με αρχή το χρονικό σημείο  $t + s$ .

Το  $s$  παριστάνει μια αυθαίρετη μετακίνηση κατά μήκος του άξονα του χρόνου είτε προς τα εμπρός είτε προς τα πίσω, δηλαδή μπορεί να είναι είτε θετικό είτε αρνητικό. Οπότε, από τη στιγμή που δεν μεταβάλλεται η συνάρτηση πιθανότητας με το χρόνο, δεν θα μεταβάλλεται ούτε η περιθωριακή συνάρτηση πιθανότητας και το ίδιο θα ισχύει και για όλες τις διμεταβλητές συναρτήσεις πιθανότητας. Όλα αυτά συνεπάγονται ότι ο μέσος και η διακύμανση της χρονοσειράς δεν μεταβάλλονται με μια αλλαγή του χρόνου, ενώ οι συνδιακυμάνσεις θα είναι συναρτήσεις μόνο της υστέρησης  $s$ .

Κάθε στάσιμη στοχαστική διαδικασία μπορεί να εκφραστεί ως γραμμικός συνδυασμός μιας ακολουθίας ασυσχέτιστων τυχαίων μεταβλητών. Ένας τέτοιος γραμμικός συνδυασμός είναι επίσης γνωστός ως γραμμικό φίλτρο (linear filter).

Στην Οικονομετρία μας ενδιαφέρει οι σειρές να είναι στάσιμες, γιατί έτσι αποφεύγονται πολλά προβλήματα, όπως αυτό της φαινομενικής παλινδρόμησης. Δυστυχώς, οι περισσότερες αν όχι όλες οι οικονομικές χρονολογικές σειρές, όπως η κατανάλωση, η ανεργία, ο δείκτης τιμών, τα κέρδη δεν έχουν τα χαρακτηριστικά στάσιμων διαδικασιών. Όμως, υπάρχει τρόπος μετατροπής μιας μη στάσιμης σειράς σε στάσιμη κι αυτό γίνεται με τη χρήση πρώτων, δεύτερων κ.τ.λ. διαφορών.

Συγκεντρωτικά να αναφέρουμε πως μια στοχαστική διαδικασία ονομάζεται στάσιμη αν ο μέσος και η διακύμανσή της δεν μεταβάλλονται διαχρονικά και η συνδιακύμανση των τιμών της σε δύο χρονικές περιόδους εξαρτάται μόνο από τις χρονικές υστερήσεις και όχι από καθαυτό το χρονικό σημείο στο οποίο υπολογίζεται (δεύτερης τάξης στασιμότητα).

Συνεπώς για μία στάσιμη χρονολογική σειρά ισχύουν οι εξής ιδιότητες:

- **Μέσος:**

$$E(Y_t) = \mu$$

- **Διακύμανση:**

$$\text{var}(Y_t) = E(Y_t - \mu)^2 = \sigma^2$$

- **Συνδιακύμανση:**

$$\gamma_k = E[(Y_t - \mu)(Y_{t+k} - \mu)]$$

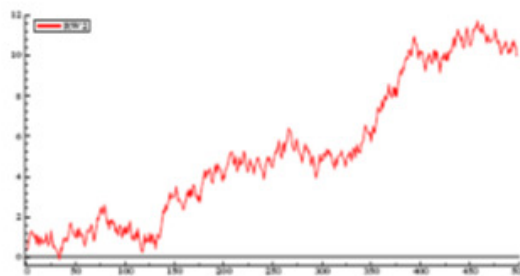
Όπου  $\gamma_k$  η συνδιακύμανση σε  $k$  χρονικές υστερήσεις δηλαδή η συνδιακύμανση μεταξύ των τιμών  $Y_t$  και  $Y_{t+k}$ .

Υπάρχουν διάφοροι τρόποι με τη βοήθεια των οποίων μπορούμε να ελέγξουμε τη στασιμότητα μιας χρονολογικής σειράς. Μερικοί από τους πιο διαδεδομένους και πιο συχνά χρησιμοποιούμενους είναι οι ακόλουθοι:

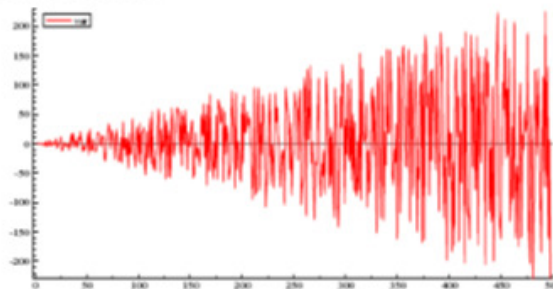
- Μελετώντας την γραφική απεικόνιση της σειράς
- Κατασκευάζοντας και μελετώντας την συνάρτηση αυτοσυσχέτισης και το αντίστοιχο κορρελόγραμμα της
- Πραγματοποιώντας έλεγχο για μοναδιαία ρίζα (Dickey-Fuller tests)
- Πραγματοποιώντας στατιστικούς ελέγχους για τον συντελεστή αυτοσυσχέτισης (Q statistic).

Ο έλεγχος για την ύπαρξη στασιμότητας ή όχι μπορεί να βασιστεί στην γραφική απεικόνιση της σειράς. Στα παρακάτω δυο γραφήματα παρατηρούμε ότι τόσο ο μέσος όσο και η διακύμανση δεν είναι σταθεροί στο χρόνο:

(a) Constant mean?



(b) Constant variance?





Σχήμα 10: Στασιμότητα ως προς τη διακύμανση λη ως προς το μέσο<sup>8</sup>

Συγκεντρωτικά, μια χρονολογική σειρά είναι στάσιμη όταν η μέση της τιμή και η διακύμανση δεν επηρεάζονται από τον χρόνο. Ενώ η συνδιακύμανση μεταξύ των τιμών της σε δύο διαδοχικά χρονικά σημεία εξαρτάται μόνο από την απόσταση ανάμεσα στα χρονικά σημεία, και όχι από τον ίδιο το χρόνο.

Μαθηματικά αυτό ισχύει όταν:

$$E(y_t) = \mu_y$$

$$V(y_t) = E(y_t - E(y_t))^2 = \sigma_y^2$$

$$\text{Cov}(y_t, y_{t-k}) = E(y_t - E(y_t))(y_{t-k} - E(y_{t-k})) = \gamma_k$$

Σε περίπτωση που μια σειρά αποδειχθεί μη στάσιμη θα πρέπει να διορθωθεί η συγκεκριμένη διαταραχή. Κάτι τέτοιο πραγματοποιείται μέσω της ακόλουθης διαδικασίας:

Πρόκειται για μη στάσιμη σειρά (χωρίς εποχικότητα)

- Δημιουργία δεδομένων πρώτης διαφοράς:

$$y'_t = y_t - y_{t-1}$$

- Άτυπος και Στατιστικός Έλεγχος Στασιμότητας για τη  $y'_t$  μέσω:  $r(k)$ , Box - Pierce, Dickey - Fuller:
- Αν  $y'_t$  Στάσιμη, Τέλος διαδικασίας
  - Αν  $y'_t$  Μη - Στάσιμη, παίρνουμε τη:

$$y''_t = y'_t - y'_{t-1}$$

<sup>8</sup> <http://slideplayer.gr/slide/1987676/>

- Και επανελέγχουμε ... έως ότου επιτύχουμε στασιμότητα.

όπου:

$$y_t'' = y_t - 2y_{t-1} + y_{t-2}$$

Σε περίπτωση που η χρονοσειρά έχει και εποχικότητα τα παραπάνω βήματα διαμορφώνονται ως εξής:

- Δημιουργία δεδομένων πρώτης διαφοράς:

$$y_t' = y_t - y_{t-12}$$

- Άτυπος και Στατιστικός Έλεγχος Στασιμότητας για τη  $y_t'$  μέσω:  $t(k)$ , Box - Pierce, Dickey - Fuller:
- Αν  $y_t'$  Στάσιμη, Τέλος διαδικασίας
  - Αν  $y_t'$  Μη - Στάσιμη, παίρνουμε τη:

$$y_t^* = y_t' - y_{t-1}'$$

- Και επανελέγχουμε ... έως ότου επιτύχουμε στασιμότητα.

όπου:

$$y_t^* = y_t - y_{t-1} - y_{t-12} + y_{t-13}$$

## 2.5. ΑΥΤΟΠΑΛΙΝΔΡΟΜΑ ΥΠΟΔΕΙΓΜΑΤΑ

Είναι γνωστό από την θεωρία των πιθανοτήτων ότι ο λόγος της συνδιακύμανσης προς το γινόμενο των τετραγωνικών ριζών των διακυμάνσεων δυο μεταβλητών είναι ο συντελεστής συσχέτισέως τους. Επίσης, ο συντελεστής συσχέτισης μας δίνει ένα μέτρο για τον βαθμό της μεταξύ τους σχέσης δυο μεταβλητών. Μάλιστα, η τιμή του συντελεστή συσχέτισης, που είναι απαλλαγμένος από τις μονάδες των μεταβλητών, δίνει

μια αρκετά πλήρη εικόνα. Και έτσι, μπορούμε να απαντήσουμε εάν η μεταξύ τους σχέση είναι ισχυρή ή ασθενής κτλ. Τέλος, για τον συντελεστή συσχέτισης ισχύει ότι

$$-1 = \rho \leq 1$$

και

- Εάν  $\rho = 1$  ή  $\rho = -1$  έχουμε την μεγίστη δυνατή συσχέτιση,
- Εάν  $\rho > 0$  υπάρχει θετική συσχέτιση, η οποία είναι τόσο πιο ισχυρή όσο πιο κοντά ο συντελεστής συσχέτισης είναι στο 1,
- Εάν  $\rho < 0$  υπάρχει αρνητική συσχέτιση, η οποία είναι τόσο πιο ισχυρή όσο πιο κοντά ο συντελεστής συσχέτισης είναι στο  $-1$ .
- Για  $\rho = 0$  δεν υπάρχει καμία συσχέτιση μεταξύ των δυο μεταβλητών.

Στην περίπτωση των χρονολογικών σειρών ο συντελεστής συσχέτισης ανάμεσα στην  $Y_t$  και στην  $Y_{t+1}$  ονομάζεται συντελεστής αυτοσυσχέτισης.

Η σημασία της συνάρτησης αυτοσυσχέτισης είναι πολύ μεγάλη, πρώτον γιατί δίνει το μέτρο της συσχέτισης των παρατηρήσεων – μετρήσεων οι οποίες απέχουν κατά το χρονικό διάστημα  $s$  και δεύτερον γιατί δείχνει τόσο το βαθμό (ένταση) όσο και το μήκος ή την χρονική διάρκεια της μνήμης της στοχαστικής διαδικασίας.

Δηλαδή, εκφράζει κατά πόσο οι μετρήσεις με χρονική απόσταση  $s$  έχουν σχέση μεταξύ τους. Με άλλα λόγια, η συνάρτηση αυτοσυσχέτισης μας δείχνει κατά πόσο το παρόν θυμάται το παρελθόν, και κατά πόσο το μέλλον θα επηρεαστεί από το παρόν.

Η γενική μορφή ενός αυτοπαλίνδρομου υποδείγματος τάξεως  $p$ , ή αλλιώς ενός AR( $p$ ) είναι:

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t$$

Η  $p$  τάξη ραναφέρεται στο μήκος της χρονικής υστέρησης. Η μορφή ενός αυτοπαλίνδρομου υποδείγματος μας παραπέμπει ουσιαστικά σε ένα υπόδειγμα παλινδρόμησης, όπου οι ερμηνευτικές μεταβλητές είναι οι τιμές της εξαρτημένης μεταβλητής με χρονική υστέρηση. Εξαιτίας αυτού του γεγονότος επικράτησε ο όρος αυτοπαλίνδρομο. Η μεταβλητή  $\varepsilon_t$  θεωρείται ότι είναι λευκός θόρυβος.

Πιο συγκεκριμένα τα Αυτοπαλίνδρομα Υποδείγματα πρώτου βαθμού έχουν την εξής μορφή:

$$\text{AR}(1) : Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon_t$$

Στα Αυτοπαλίνδρομα Υποδείγματα πρώτου βαθμού η χρονολογική σειρά παλινδρομείται σε μια χρονική υστέρηση της σειράς.

Οι βασικές ιδιότητες αυτών των υποδειγμάτων είναι οι ακόλουθες:

**Μέσος:**

$$E[Y_t] = \mu = \frac{\beta_0}{(1 - \beta_1)}$$

**Διακύμανση:**

$$\text{Var}[Y_t] = \frac{\sigma_\varepsilon^2}{1 - \beta_1^2}$$

**Συνδιακύμανση:**

$$\text{Cov}(Y_t, Y_{t-s}) = \frac{\beta_1^s \sigma_\varepsilon^2}{1 - \beta_1^2}$$

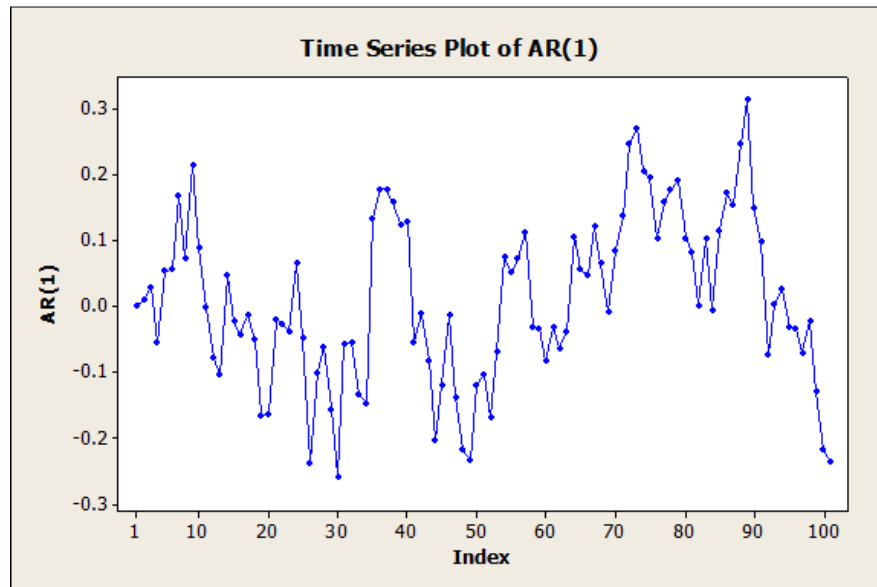
**Συσχέτιση:**

$$\text{Corr}(Y_t, Y_{t-s}) = \beta_1^s$$

Με τη χρήση αυτού του υποδείγματος οι προβλέψεις γίνονται με βάση το ακόλουθο υπόδειγμα:

$$E[Y_{T+1} | Y_T] = \beta_0 + \beta_1 Y_T + E[\varepsilon_{T+1} | Y_T] = \beta_0 + \beta_1 Y_T$$

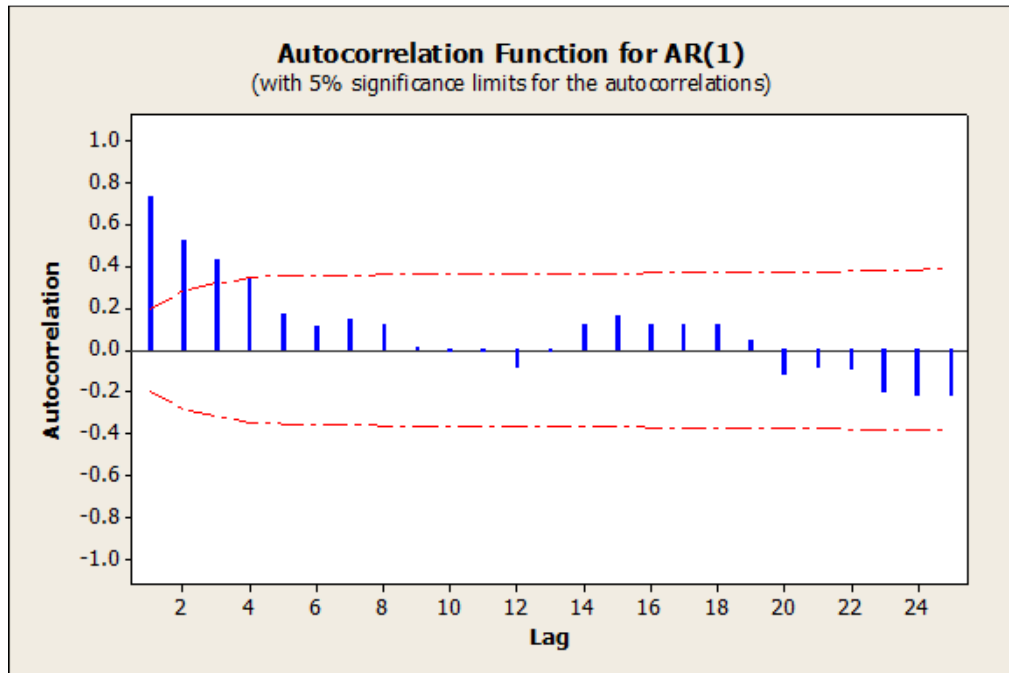
Η γραφική απεικόνιση ενός τέτοιου υποδείγματος θα μοιάζει όπως η ακόλουθη γραφική παράσταση:



Σχήμα 11: Αυτοσυσχέτιση 1<sup>ης</sup> τάξης<sup>9</sup>

Στην περίπτωση αυτή η συνάρτηση αυτοσυσχέτισης θα έχει την εξής μορφή:

<sup>9</sup> <http://pubs.sciepub.com/ajams/2/1/4/>



Σχήμα 12: Αυτοπαλίνδρομο υπόδειγμα AR(1)<sup>10</sup>

Για να είναι στάσιμη μία χρονολογική σειρά που ακολουθεί ένα υπόδειγμα AR(1) πρέπει:

$$|\beta_1| < 1$$

## 2.6. ΥΠΟΔΕΙΓΜΑΤΑ ΚΙΝΗΤΟΥ ΜΕΣΟΥ

Οι διαδικασίες κινητού μέσου είναι χρήσιμες για περιγραφή φαινομένων στα οποία τα γεγονότα παράγουν ένα άμεσο αποτέλεσμα, η επίδραση του οποίου όμως δεν σταματά εκεί αλλά διαρκεί, αν και το ίδιο το γεγονός παύει να υφίσταται. Συνήθως επηρεάζει λιγότερο και για μικρό χρονικό διάστημα τις επόμενες χρονικές στιγμές. Για παράδειγμα μια απεργία επηρεάζει την οικονομία όχι μόνο την στιγμή που πραγματοποιείται αλλά και τους επομένους μήνες αν και έχει λήξει.

<sup>10</sup> <https://onlinecourses.science.psu.edu/stat510/node/60>

Η γενική μορφή ενός υποδείγματος κινητού μέσου τάξης  $q$  ή  $MA(q)$  είναι:

$$Y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Η τάξη  $q$  αναφέρεται στο μήκος της υστέρησης της μεταβλητής, για την οποία υποθέτουμε ότι είναι λευκός θόρυβος. Ο όρος κινητός μέσος αναφέρεται στο γεγονός ότι η  $Y_t$  εμφανίζεται ως ένα σταθμισμένο άθροισμα των τιμών της  $\varepsilon_t$ .

Πιο συγκεκριμένα αναφορικά με το υπόδειγμα κινητού μέσου πρώτης τάξης (πρώτου βαθμού  $MA(1)$ ) ισχύουν τα ακόλουθα:

Έστω  $\{e_t\}$  λευκός θόρυβος. Ένα υπόδειγμα κινητών μέσων μπορεί να εκφραστεί ως εξής:

$$y_t = \mu + e_t + \theta_1 e_{t-1}$$

Δηλαδή οι τιμές της χρονολογικής σειράς είναι συνάρτηση του διαταραχτικού όρου της τρέχουσας περιόδου και του διαταραχτικού όρου της προηγούμενης περιόδου.

Οι ιδιότητες αυτού είναι οι ακόλουθες:

**Μέσος:**

$$E y_t = E(\mu + e_t + \theta_1 e_{t-1}) = \mu$$

**Διακύμανση:**

$$Var Y_t = (1 + \theta_1^2) \sigma^2$$

**Συνδιακύμανση:**

$$Cov(Y_t, Y_{t-1}) = -\theta_1 \sigma^2$$

Η χρονολογική σειρά που ακολουθεί ένα  $MA$  υπόδειγμα είναι πάντα στάσιμη.

## 2.7. ΤΑΣΗ ΚΑΙ ΕΠΟΧΙΚΟΤΗΤΑ

Η μακροχρόνια τάση είναι από τις περισσότερο εύκολα αντιμετωπίσιμες περιπτώσεις γιατί η τάση εκφράζει την χρονολογική σειρά για μια εκτεταμένη περίοδο. Οι χρονολογικές σειρές που παρουσιάζουν μακροχρόνια τάση είναι αυτές που έχουν μελετηθεί περισσότερο. Επίσης, στις χρονολογικές σειρές με μακροχρόνια τάση είναι εύκολη η κατανόηση της ιστορίας της μεταβλητής με συνέπεια την πρόβλεψη μελλοντικών τιμών. Ορισμένες μέθοδοι προσδιορισμού της μακροχρόνιας τάσης είναι η μέθοδος των δυο μέσων σημείων, η μέθοδος των κινητών μέσων, η μέθοδος της ευθείας ελαχίστων τετραγώνων, η μέθοδος της καμπύλης ελαχίστων τετραγώνων και άλλα.

Η κυκλική κύμανση που παρουσιάζουν περιπτώσεις χρονολογικών σειρών έγκειται στο γεγονός ότι παρατηρούνται αυξομειώσεις της τιμής της μεταβλητής γύρω από γραμμή τάσης σε μια μακροχρόνια περίοδο. Στην πράξη τα σημεία της χρονολογικής σειράς για μια χρονική σειρά παρατηρήσεων βρίσκονται κάτω από την γραμμή τάσης και στην συνέχεια για άλλη χρονική σειρά τιμών πάνω από την γραμμή τάσης. Ο χρόνος για να έχουμε μια κυκλική αυξομείωση δεν είναι σταθερός.

Στην πράξη οι κυκλικές αυξομειώσεις είναι οι πλέον δύσκολες να αντιμετωπιστούν γιατί η κυκλική κίνηση δεν ακολουθεί κανένα κανονικό μοντέλο αλλά κινείται απρόβλεπτα.

Αντίθετα, από τις χρονολογικές σειρές που παρουσιάζουν κυκλικές μεταβολές και είναι δύσκολα αντιμετωπίσιμες, οι χρονολογικές σειρές που παρουσιάζουν περιοδικές μεταβολές είναι πολύ χρήσιμες γιατί ακολουθούν κανονικό μοντέλο και έτσι μπορούν να δώσουν αξιόπιστες προβλέψεις για το μέλλον.

Τέλος, υπάρχουν και χρονολογικές σειρές που παρουσιάζουν ακανόνιστες μεταβολές, μεταβολές που άλλοτε είναι μικρές, άλλοτε μεγάλες, θετικές ή αρνητικές χωρίς καμία κανονικότητα. Οι μεταβολές αυτές διακρίνονται σε συμπτωματικές, οφειλόμενες σε απρόβλεπτα γεγονότα και τυχαίες.



Αναφορικά με τα υποδείγματα μέτρησης της τάσης μίας σειράς να αναφέρουμε πως έχουν ως εξαρτημένη μεταβλητή τις τιμές της χρονολογικής σειράς και ως ανεξάρτητη μεταβλητή την χρονική περίοδο.

Οι πιο συνήθεις εξειδικεύσεις είναι οι ακόλουθες:

- **Γραμμική τάση**

$$Y_t = a + b t + e_t$$

- **Λογαριθμική Τάση**

$$Y_t = a + b \ln(t) + e_t$$

- **Εκθετική Τάση**

$$Y_t = \exp\{a + b t + e_t\}$$

- **Πολυωνομική Τάση**

$$Y_t = a + b_1 t + b_2 t^2 + e_t$$

Η ύπαρξη εποχικότητας αναγνωρίζεται ως ένα φαινόμενο επαναλαμβανόμενων σε μέγεθος τιμών σε συγκεκριμένη χρονική στιγμή.

- Παράδειγμα Α: Μέγιστες και ελάχιστες τιμές αεροπορικών επιβατών.
- Παράδειγμα Β: Ελάχιστες τιμές θερμοκρασίας στο Παρίσι.

## **2.8. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ – ΧΡΗΣΙΜΟΤΗΤΑ – ΔΙΑΣΤΡΩΜΑΤΙΚΩΝ ΣΤΟΙΧΕΙΩΝ**

Τα δεδομένα χρονολογικών σειρών που αναλύθηκαν στις προηγούμενες ενότητες χρησιμοποιούνται κατά κύριο λόγο για τη μελέτη μιας επιχείρησης σε βάθος χρόνου. Αντίθετα τα διαστρωματικά στοιχεία χρησιμοποιούνται με σκοπό να συγκρίνουν στην ίδια μονάδα χρόνου, («την ίδια χρονική περίοδο») περισσότερες από μια μονάδες (π.χ. επιχειρήσεις, μετοχές, αγαθά, υπηρεσίες, κ.λπ.).

Η διαστρωματική ανάλυση συχνά χρησιμοποιείται σε χρηματοοικονομικές αναλύσεις και αναλύσεις λογιστικών καταστάσεων όπου ως διαστρωματική ανάλυση (comparative analysis or cross-sectional analysis), ορίζεται η σύγκριση των χρηματοοικονομικών στοιχείων της επιχείρησης με τα αντίστοιχα παρόμοιων επιχειρήσεων ή τη μέση τιμή του κλάδου στον οποίο ανήκει κατά τη διάρκεια του ίδιου χρονικού διαστήματος.

Στην περίπτωση ανάλυσης δεδομένων μιας περιόδου για περισσότερες από μια επιχειρήσεις, έχουν ισχύ οι μεθοδολογίες της διαστρωματικής ανάλυσης. Έτσι ο αναλυτής, έχει στη διάθεσή του μια σημαντική ποικιλία μεθοδολογιών ποσοτικής ανάλυσης, οι οποίες του επιτρέπουν να πραγματοποιεί εκτιμήσεις με ανεκτές αποκλίσεις. Είναι αυτονόητο ότι δεν υπάρχουν περιορισμοί ως προς τη μορφή των μεταβλητών που χρησιμοποιούνται στις αναλύσεις χρονολογικών σειρών ή διαστρωματικών δεδομένων. Έτσι, αυτές είναι δυνατό να έχουν τη μορφή απόλυτων μεγεθών, ποσοστών μεταβολής κλπ. Πάντως, στην πράξη, χρησιμοποιούνται κατά βάση δείκτες, οι οποίοι προκύπτουν από τη συσχέτιση δυο μεμονωμένων μεγεθών ή δυο ομάδων μεγεθών της επιχείρησης, του κλάδου κλπ.

Στα πλαίσια της ανάλυσης της οικονομικής κατάστασης των επιχειρήσεων, η διαστρωματική ανάλυση περιλαμβάνει την ποσοτική ανάλυση στοιχείων ενός αριθμού οικονομικών μονάδων, που συνήθως είναι ταξινομημένες στον ίδιο κλάδο ή υποκλάδο.

Οι στατιστικές / οικονομετρικές μεθοδολογίες που εφαρμόζονται, διαφέρουν ανάλογα με τη μορφή και την έκταση των διαθέσιμων στοιχείων και το είδος των ζητούμενων πληροφοριών. Συχνότερα, πάντως, χρησιμοποιείται η μέθοδος των ελαχίστων τετραγώνων (ordinary least squares) και υπολογίζονται τα βασικά στατιστικά

μέτρα περιγραφής των υπό εξέταση κατανομών (μέσος όρος, τυπική απόκλιση, διάμεσος, δεκατημόρια, συντελεστές ασυμμετρίας και κύρτωσης).

Για να αντιμετωπιστούν οι διαφορές κλίμακας που υπάρχουν μεταξύ των επιχειρήσεων του κάθε φορά εξεταζόμενου δείγματος, τα δεδομένα εκφράζονται σε σχετικούς όρους, όπως δείκτες, εκατοστιαίες αναλογίες κλπ. Πάντως, δεν αποκλείεται και η χρήση απόλυτων αριθμών, όταν η φύση των υπό ανάλυση μεγεθών και η χρησιμοποιούμενη μεθοδολογία το επιτρέπουν.

Οι βασικές εφαρμογές της διαστρωματικής ανάλυσης περιλαμβάνουν την αποτίμηση μιας εταιρίας, την αξιολόγηση των επιδόσεων του management και την εκτίμηση της πιθανότητας χρεοκοπίας της.

Ένα μεγάλο πρόβλημα της διαστρωματικής ανάλυσης συνδέεται με την διαθεσιμότητα των αναγκαίων δεδομένων. Οι συνηθέστερες δυσκολίες που παρατηρούνται στον τομέα αυτό έχουν ως εξής:

- Μη διαθέσιμα δεδομένα
- Ανομοιομορφία στις λογιστικές στρατηγικές
- Δεδομένα που αναφέρονται σε διαφορετικές χρονικές περιόδους.

## **2.9. ΠΕΛΙΟ ΕΦΑΡΜΟΓΗΣ**

Τα διαστρωματικά στοιχεία αφορούν μεγάλο αριθμό στατιστικών παρατηρήσεων για πολλές χώρες για συγκεκριμένα χρονικά διαστήματα (πχ. τριαντακονταετία). Στην οικονομική μεγέθυνση οι παρατηρήσεις αυτές αφορούν συνήθως μερικές παρατηρήσεις (1-5) για μεγάλο αριθμό οικονομιών. Οι σειρές αυτές μπορεί να είναι είτε παρατηρήσεις για μία συγκεκριμένη περίοδο για πολλές οικονομίες (πχ. εισόδημα και κεφάλαιο μιας συγκεκριμένης χρονιάς για όλες τις οικονομίες) είτε να αφορούν το μέσο όρο μιας ή περισσότερων περιόδων για πολλές οικονομίες. Η τελευταία προσέγγιση είναι ιδιαίτερα διαδεδομένη, γιατί η χρήση μέσων όρων από μεγάλα διαστήματα επιτρέπει την εξαγωγή συμπερασμάτων για τη μακροχρόνια ισορροπία των οικονομιών. Με τα διαστρωματικά

στοιχεία μπορούν λοιπόν να εξαχθούν χρήσιμα συμπεράσματα για την κατάσταση ισορροπίας από πολλές οικονομίες, αλλά όμως υπάρχει ο κίνδυνος να χρησιμοποιηθούν στοιχεία που προέρχονται από ανομοιογενείς πηγές.

Η βασική προσέγγιση στην εμπειρική εκτίμηση των επιδράσεων στην οικονομική μεγέθυνση προέρχεται από τις διαστρωματικές παλινδρομήσεις, οι οποίες χρησιμοποιούν μέσους όρους από μεγάλα χρονικά διαστήματα για πολλές οικονομίες, με στόχο την ανάλυση των παραγόντων, που επιδρούν στην οικονομική μεγέθυνση σε μακροχρόνιο επίπεδο. Η προσέγγιση αυτή, σε συνδυασμό με την μεγαλύτερη ευρύτητα των διαθέσιμων στοιχείων μέσω της βάσης δεδομένων Penn World Tables (βλ. Κεφάλαιο 4) και των στοιχείων της Διεθνούς Τράπεζας, έχει οδηγήσει σε εκτιμήσεις πολυμεταβλητών εξισώσεων με στοιχεία από μεγάλο αριθμό χωρών. Οι εξισώσεις αυτές είναι γνωστές στη βιβλιογραφία της οικονομικής μεγέθυνσης ως 'παλινδρομήσεις τύπου Barro', διότι πρωτοπαρουσιάστηκαν από τον Barro (1991).

Το κύριο χαρακτηριστικό αυτών των εξισώσεων είναι ότι χρησιμοποιούν μια σειρά από υποψήφιες μεταβλητές, οι οποίες θεωρούνται ότι είναι πιθανόν να επηρεάζουν την οικονομική μεγέθυνση. Οι μεταβλητές επιλέγονται είτε με βάση τη θεωρία (π.χ. επενδύσεις) είτε επειδή η συσχέτισή τους με την οικονομική μεγέθυνση θεωρείται προφανής (π.χ. θεσμικό καθεστώς). Οι μεταβλητές που συνήθως περιλαμβάνονται στις σχετικές παλινδρομήσεις είναι η σχετική αρχική κατάσταση της οικονομίας (που προσεγγίζεται από το αρχικό εισόδημα), οι επενδύσεις (ως ποσοστό του ΑΕΠ), το ανθρώπινο κεφάλαιο (που συνήθως προσεγγίζεται από το βαθμό συμμετοχής στην εκπαίδευση), οι δημόσιες δαπάνες (που υπολογίζονται από τη δημόσια κατανάλωση ως ποσοστό του ΑΕΠ), και οι διεθνείς συναλλαγές (εισαγωγές και εξαγωγές ως ποσοστό του ΑΕΠ, ανταγωνιστικότητα). Ακόμα, συνυπολογίζεται η επίδραση άλλων παραγόντων, η επίδραση των οποίων συνήθως δεν αναλύεται σε θεωρητικά υποδείγματα, αλλά είναι προφανής από την οικονομική πραγματικότητα (καθεστώς της αγοράς, πληθωρισμός, θεσμικό περιβάλλον, πολιτική σταθερότητα), Τέλος, λαμβάνονται πάντα υπόψη πιθανά οικονομετρικά προβλήματα (σφάλματα στατιστικής μέτρησης που παρατηρούνται στις υποανάπτυκτες χώρες, ετερογένεια των στοιχείων από διαφορετικές χώρες, κατεύθυνση της αιτιότητας, πιθανότητα μη-γραμμικής σχέσης μεταξύ των μεταβλητών).

Ένας μεγάλος αριθμός αναλύσεων έχει γίνει με χρήση διαστρωματικών στατιστικών στοιχείων (στο επίπεδο της παραγωγικής μονάδας). Η ομάδα αυτή που χρησιμοποιεί διαστρωματικά στοιχεία στο επίπεδο της παραγωγικής μονάδας περιλαμβάνει αναλύσεις που έχουν αποκλειστικά χρησιμοποιήσει τη συνάρτηση κέρδους. Ο λόγος είναι απλός. Στο επίπεδο της παραγωγικής μονάδας η μεγιστοποίηση του κέρδους, μάλλον παρά η ελαχιστοποίηση του κόστους, αποτελεί τον κανόνα συμπεριφοράς. Επομένως, το επίπεδο παραγωγής ορίζεται ενδογενώς.

Στις αναλύσεις αυτές έχουν χρησιμοποιηθεί διάφορες συναρτησιακές μορφές. Οι πρώτες εφαρμογές είχαν χρησιμοποιήσει μια συνάρτηση κέρδους της μορφής Cobb-Douglas.

Όπως ήδη αναφέρθηκε η αξιολόγηση επενδύσεων και η χρηματοοικονομική ανάλυση μιας επιχείρησης είναι μάλλον ο κλάδος στον οποίο εφαρμόζονται περισσότερο τα διαστρωματικά στοιχεία έτσι ώστε να διευκολύνεται η σύγκριση καταστάσεων και η εξαγωγή συμπερασμάτων, συντελώντας κατά αυτό τον τρόπο στην ορθότερη λήψη αποφάσεων.

## **2.10. ΤΟ ΠΡΟΒΛΗΜΑ ΤΗΣ ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑΣ**

Όταν τα στατιστικά δεδομένα είναι υπό διαστρωματική μορφή, τότε υπάρχει κίνδυνος εμφάνισης ετεροσκεδαστικότητας. Σε αυτή την περίπτωση, η εκτίμηση των συντελεστών παλινδρόμησης και ο έλεγχος υποθέσεων καθίστανται προβληματικοί, επειδή η διακύμανση των σφαλμάτων της παλινδρόμησης δεν είναι σταθερή, αλλά μεταβάλλεται.

Αυτό οφείλεται στο γεγονός ότι η διακύμανση των σφαλμάτων συνδέεται με διάφορους τρόπους με την τιμή της ερμηνευτικής μεταβλητής. Κατά αυτό τον τρόπο οι εκτιμήσεις των συντελεστών της παλινδρόμησης ασυμπτωτικά δεν είναι αποτελεσματικές. Για τον εντοπισμό της ετεροσκεδαστικότητας έχουν αναπτυχθεί διάφοροι έλεγχοι οι οποίοι και αναλύονται εκτενώς σε προηγούμενη ενότητα. Να σημειωθεί πως οι σημαντικότεροι από αυτούς είναι το κριτήριο του White, του Glejser

και του Bartlett. Επιπλέον να αναφέρουμε πως δεν αρκεί να εντοπίσουμε την ύπαρξη ή μη ετεροσκεδαστικότητας, πρέπει και να την επιλύσουμε/ διορθώσουμε.

Η συνήθης πρακτική στην εν λόγω περίπτωση είναι η μελέτη της συμπεριφοράς των καταλοίπων και των εμπλεκόμενων μεταβλητών μέσω των διαγραμμάτων διασποράς αυτών (scatter plots).

## **2.11. ΔΙΑΣΤΡΩΜΑΤΙΚΑ ΣΤΟΙΧΕΙΑ vs ΧΡΟΝΟΛΟΓΙΚΕΣ ΣΕΙΡΕΣ**

Συμπερασματικά θα μπορούσαμε να αναφέρουμε πως κάθε μια από τις δυο αναλύσεις – προσεγγίσεις έχει τα δικά της οφέλη και πλεονεκτήματα αφενός αλλά και τους δικούς της περιορισμούς και δυσκολίες από την άλλη. Η επιλογή ανάμεσα στις δυο τεχνικές γίνεται ανάλογα με τον ερευνητή και τους στόχους που αυτός έχει θέσει.

Ο συνδυασμός των δυο, δηλαδή η ανάλυση δεδομένων panel αν και είναι χρονοβόρα και κοστίζει είναι αυτή που οδηγεί στα πιο πλήρη αποτελέσματα δίνοντας τη δυνατότητα στους μελετητές να εξάγουν ασφαλή συμπεράσματα και παράλληλα να προβούν στις κατάλληλες προβλέψεις για τα μεγέθη τα οποία μελετούν.

### ΚΕΦΑΛΑΙΟ 3: «ΑΝΑΛΥΣΗ ΔΕΔΟΜΕΝΩΝ ΠΑΝΕΛ»

Σε συνέχεια των προηγούμενων δυο κεφαλαίων στο παρόν αναλύεται η έννοια των δεδομένων panel, πως αυτά χρησιμοποιούνται, από ποιες υποθέσεις διέπονται και ποια μοντέλα τα μαθηματικοποιούν.

Η ανάλυση δεδομένων πάνελ (panel data) αποτελεί μια μέθοδο που συνδυάζει διαστρωματικά στοιχεία με χρονολογικές σειρές, μέσω της οποίας εξετάζονται τόσο η χρονική όσο και η διαστρωματική παράμετρος. Η μέθοδος αυτή στηρίζεται στη συλλογή δεδομένων κατόπιν παρακολούθησης –παρατήρησης διαχρονικά (ανά εβδομάδα, μήνα, έτος κλπ) συγκεκριμένων μονάδων.

Παρόλο που η ύπαρξη panel data έδωσε την δυνατότητα στην εμπειρική οικονομική έρευνα να ασχοληθεί με νέους τρόπους ανάλυσης της συμπεριφοράς των οικονομικών μονάδων (παραγωγικών ή καταναλωτικών), από την άλλη πλευρά έγειρε σημαντικά προβλήματα ως προς την χρήση των κατάλληλων οικονομετρικών υποδειγμάτων. Παρόλο που η ανάπτυξη οικονομετρικών υποδειγμάτων για την ανάλυση διαστρωματικών στοιχείων ξεκινάει από τις αρχές της δεκαετίας του 60, τα τελευταία 15 χρόνια παρουσιάστηκε μία έντονη μεθοδολογική έρευνα που αποσκοπούσε στην προσαρμογή και ανάπτυξη των οικονομετρικών υποδειγμάτων για την αντιμετώπιση προβλημάτων που αφορούν την εμπειρική εξειδίκευση, οικονομετρική εκτίμηση και στατιστικό έλεγχο οικονομικών υποδείγματος ανάλυσης της συμπεριφοράς των οικονομικών μονάδων. Ιδιαίτερα σημαντική προς την κατεύθυνση αυτή ήταν η συμβολή των J. Heckman και D. McFadden οι οποίοι για τον λόγο αυτό τιμήθηκαν με το Νόμπελ Οικονομίας το 2000.

Για να χρησιμοποιήσουμε δεδομένα πάνελ σε μια οικονομετρική μελέτη είναι σημαντικό να γνωρίζουμε πως πρέπει να γίνεται η καταχώρηση των δεδομένων. Πρέπει να φροντίσουμε ώστε η διάταξη των δεδομένων να γίνει με τέτοιο τρόπο ώστε να μπορεί να γίνει εύκολα σύνδεση των διαφορετικών χρονικών περιόδων για την ίδια μονάδα διαστρωμάτωσης (άτομο, επιχείρηση, πόλη, και ούτω καθεξής).

### 3.1. ΣΤΟΧΟΣ – ΣΚΟΠΟΣ - ΧΡΗΣΙΜΟΤΗΤΑ

Ο συνδυασμός διαχρονικών και διαστρωματικών δεδομένων προσφέρει στον ερευνητή μεγαλύτερη ευελιξία σε σχέση με τις παραδοσιακές μορφές ανάλυσης.

Ένα σύνολο από “panel” στοιχεία περιέχει παρατηρήσεις από πολλαπλές «οντότητες», όπου κάθε οντότητα παρατηρείται σε δύο ή περισσότερα χρονικά σημεία.

Μερικά παραδείγματα τέτοιων δεδομένων είναι τα ακόλουθα:

- Στοιχεία για 420 σχολικές μονάδες στην Καλιφόρνια το 2012 και ξανά το 2013, για 840 παρατηρήσεις, συνολικά.
- Στοιχεία για 50 πολιτείες των ΗΠΑ, κάθε πολιτεία «παρατηρείται» σε 3 έτη, για 150 παρατηρήσεις, συνολικά.
- Στοιχεία για 1000 άτομα, σε τέσσερις διαφορετικούς μήνες, για 4000 παρατηρήσεις, συνολικά.

Οι συμβολισμοί για τα στοιχεία “panel” είναι συνήθως οι ακόλουθοι:

Ένας διπλός δείκτης διαχωρίζει τις οντότητες (στρώματα) και τις χρονικές περιόδους (έτη)

$i$  = οντότητα (στρώμα),  $n$  = πλήθος στρωμάτων έτσι:

$$i = 1, \dots, n$$

$t$  = χρονική περίοδος (έτος),  $T$  = πλήθος χρονικών περιόδων έτσι:

$$t = 1, \dots, T$$

Έστω ότι έχουμε μία ερμηνευτική μεταβλητή. Τα στοιχεία είναι:

$$(X_{it}, Y_{it}), i = 1, \dots, n, t = 1, \dots, T$$

Στοιχεία “panel” με  $k$  ερμηνευτικές μεταβλητές:



$$(X_{1it}, X_{2it}, \dots, X_{kit}, Y_{it}), i = 1, \dots, n, t = 1, \dots, T$$

Όπου:

- $n$  = πλήθος στρωμάτων
- $T$  = πλήθος χρονικών περιόδων.

Να αναφέρουμε στο σημείο αυτό ότι μια εναλλακτική ονομασία για τα στοιχεία “panel” είναι τα στοιχεία “longitudinal”. Δυο σημαντικές κατηγορίες δεδομένων είναι:

- τα **“balanced panel”**: καμία παραλειπόμενη παρατήρηση και
- τα **“unbalanced panel”**: κάποιες οντότητες (στρώματα) δεν παρατηρούνται σε κάποια χρονικά σημεία (έτη)

### 3.1.1. ΠΛΕΟΝΕΚΤΗΜΑΤΑ & ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΤΩΝ ΔΕΔΟΜΕΝΩΝ PANEL

Τα κοινά χαρακτηριστικά των υποδειγμάτων παλινδρόμησης με panel data είναι:

- Το  $N$  είναι σχετικά μεγάλο (για τα Random Effects υποδείγματα είναι ακόμη μεγαλύτερο από ότι για τα Fixed Effects)
- Το  $T$  αναφέρεται συνήθως σε μια σύντομη χρονική περίοδο ( $T \geq 2$ ).

Ο Baltagi παραθέτει μια σειρά από κίνητρα αναφορικά με τη χρησιμοποίηση των πάνελ δεδομένων έναντι των διαστρωματικών δεδομένων ή των δεδομένων χρονοσειρών στην οικονομετρική ανάλυση. Στις παραγράφους που ακολουθούν αναλύονται αυτά τα πλεονεκτήματα αλλά και ορισμένα από τα μειονεκτήματα της χρήσης των δεδομένων πάνελ. Πιο συγκεκριμένα, τα πλεονεκτήματα αυτών είναι τα ακόλουθα:

Καθώς τα πάνελ δεδομένα μελετούν άτομα, εταιρείες, οικογένειες, χώρες, κ.λπ., κατά την πάροδο του χρόνου, είναι απολύτως βέβαιο ότι θα υπάρχει ετερογένεια σε αυτές τις μονάδες – οντότητες. Οι τεχνικές εκτίμησης των πάνελ δεδομένων μπορούν να λάβουν υπόψη τους αυτή την ετερογένεια στα δεδομένα.

Αν μελετάμε κάποιο δυναμικό φαινόμενο, δηλαδή ένα φαινόμενο που μεταβάλλεται προϊόντος του χρόνου, π.χ. την εξαμηνιαία κερδοφορία ενός συνόλου επιχειρήσεων του κλάδου τροφίμων για μια περίοδο 5 ετών, τα panel data είναι τα πιο κατάλληλα, καθώς τα μεν διαστρωματικά στοιχεία (cross section data) δεν μπορούν να εκφράσουν τέτοιου είδους δυναμικές σχέσεις, τα δε στοιχεία χρονολογικών σειρών (time series data) εκφράζουν δυναμικές σχέσεις, αλλά οι εκτιμήσεις που παράγουν δεν είναι ιδιαίτερα ακριβείς, λόγω της ύπαρξης πολυσυγγραμμικότητας.

Η ανάλυση panel δεδομένων παρέχει πιο αξιόπιστη εκτίμηση των παραμέτρων του υποδείγματος, μιας και τα δείγματα που συνδυάζουν διαχρονικά και διαστρωματικά δεδομένα περιέχουν περισσότερους βαθμούς ελευθερίας και μεγαλύτερη δειγματική μεταβλητότητα από τα διαστρωματικά δεδομένα τα οποία μπορούν να θεωρηθούν ως panel με  $T = 1$  ή τα διαχρονικά στοιχεία τα οποία μπορούν να θεωρηθούν ως πάνελ με  $N=1$ .

Συνδυάζοντας χρονοσειρές διαστρωματικών παρατηρήσεων, τα πάνελ δεδομένα προσδίδουν περισσότερα πληροφοριακά στοιχεία, μεγαλύτερη μεταβλητότητα, λιγότερη συγγραμμικότητα μεταξύ των μεταβλητών περισσότερους βαθμούς ελευθερίας καθώς και περισσότερη αποτελεσματικότητα.

Επιπλέον η ανάλυση αυτή επιτρέπει την καλύτερη προσέγγιση της πολυπλοκότητας της ανθρώπινης συμπεριφοράς μέσω του ελέγχου της επίδρασης παραλειφθέντων (omitted) μεταβλητών. Ο συνδυασμός διαστρωματικών και διαχρονικών δεδομένων περιέχει πληροφορίες τόσο για την διαχρονική δυναμική όσο και για τα μοναδικά χαρακτηριστικά των διαστρωματικών μονάδων επιτρέποντας με αυτό τον τρόπο τον έλεγχο της επίδρασης παραλειπόμενων ή μη παρατηρήσιμων επιδράσεων.

Τα panel data παρέχουν εκτιμήσεις αυξημένης ακρίβειας, καθώς το πλήθος των συνολικών παρατηρήσεων που χρησιμοποιούν είναι υπερδιπλάσιο σε σχέση με το πλήθος των παρατηρήσεων που χρησιμοποιείται τόσο κατά την εκτίμηση με διαστρωματικά στοιχεία, όσο και κατά την εκτίμηση με στοιχεία χρονολογικών σειρών. Αν για παράδειγμα, έχουμε έναν δείγμα 50 χωρών ( $N = 50$ ) και θέλουμε να ελέγξουμε τις

τιμές των κρατικών ομολόγων για μια περίοδο 10ετών ( $T = 50$ ), το πλήθος των παρατηρήσεων που θα χρησιμοποιήσουμε για την ανάλυση με panel data θα είναι:

$$N \times T = 500.$$

Παρέχουν τη δυνατότητα ελέγχου των σταθερών χαρακτηριστικών, δηλαδή των χαρακτηριστικών που μεταβάλλονται μεταξύ των  $i$ ' s, αλλά παραμένουν διαχρονικά σταθερά. Συνεπώς, τα panel data λαμβάνουν υπόψη τους μεγαλύτερο ετερογένειας.

Σχηματικά, λοιπόν, έχουμε: μείωση της πολυσυγγραμμικότητας, αύξηση της ακρίβειας των εκτιμήσεων, αύξηση της μεταβλητότητας (διακύμανσης) των τιμών των ερευνητικών μεταβλητών  $x_{it}$  για κάθε δοσμένη χρονική περίοδο  $t$ .

Μας δίδουν τη δυνατότητα να υποδειματοποιήσουμε μια διαχρονική συμπεριφορά που χαρακτηρίζει κάθε ξεχωριστή οντότητα  $i$ , αποφεύγοντας έτσι το σφάλμα προσθετικότητας: μια διαχρονική συμπεριφορά που χαρακτηρίζει κάποιο ή κάποια από τα  $i$ ' s, δε σημαίνει ότι θα πρέπει να χαρακτηρίζει και το σύνολο των  $i$ ' s.

Μας επιτρέπουν να διορθώσουμε τη συνδυασμένη επίδραση (αποτέλεσμα) της μεροληψίας από παραλειπόμενες μεταβλητές στο υπόδειγμα, δηλαδή μεταβλητών που έχουν παραληφθεί από το υπόδειγμα και οι οποίες διαφέρουν μεταξύ των  $i$ ' s, αλλά είναι διαχρονικά σταθερές. Το  $a_i$  προσομοιώνει τη συνδυασμένη αυτή επίδραση των παραλειπόμενων μεταβλητών στο υπόδειγμα.

Σε συνέχεια των παραπάνω να αναφέρουμε πως τα μειονεκτήματα των πάνελ δεδομένων είναι τα ακόλουθα:

- **Φθορά:** μια από τις επιχειρήσεις που λαμβάνουν μέρος στο πείραμα, δηλαδή αποτελεί μέρος των στοιχείων του δείγματος χρεοκοπεί προτού να ολοκληρωθεί το πείραμα και να αναγκάζεται να αποχωρήσει από αυτό. Έχουμε συνεπώς απώλεια πληροφόρησης και αποτελεσματικότητας.
- **Μη τυχαίο δείγμα:** αν η φθορά έχει συστηματικό χαρακτήρα, π.χ. αν από τον κλάδο τροφίμων που εξετάζουμε αποχωρήσουν για κάποιο λόγο όλες οι εταιρείες που κατασκευάζουν ζυμαρικά, έχουμε εμφάνιση μεροληψίας επιλογής και, άρα,

το σύνολο των panel data χάνει ένα σημαντικό μέρος της αντιπροσωπευτικότητας του, έτσι ώστε να μπορεί να χαρακτηριστεί ως μη τυχαίο.

Για την εφαρμογή πάνελ δεδομένων σε μια οικονομετρική ανάλυση είναι απαραίτητο να έχουν τα στοιχεία συγκεκριμένη δομή, έτσι ώστε η μονάδα διαστρωμάτωσης να συνδέεται με τη χρονική μονάδα στην οποία αναφέρεται. Είναι σύνηθες σε δεδομένα πάνελ ο αριθμός των διαστρωματικών στοιχείων να είναι μεγάλος συγκριτικά με τον αριθμό των χρονικών περιόδων κι εστιάζεται το ενδιαφέρον μας στην ετερογένεια που οφείλεται στις επιδράσεις μη παρατηρούμενων μεταβλητών.

Η ύπαρξη panel data υποδηλώνει από μόνη της ότι οι οικονομικές μονάδες, οι περιφέρειες ή οι χώρες έχουν διαφορετικά εγγενή χαρακτηριστικά. Η χρήση απλών χρονολογικών σειρών ή διαστρωματικών στοιχείων αγνοεί αυτή τη διαφορετικότητα μεταξύ τους γεγονός που μπορεί να οδηγήσει σε μεροληπτικές εκτιμήσεις (Moulton, 1986; 1987). Παραδείγματος χάρη οι Baltagi και Levin (1992) προσπάθησαν να διερευνήσουν την κατανάλωση τσιγάρων σε 46 αμερικανικές πολιτείες κατά την περίοδο 1963-1988. Η κατανάλωση τσιγάρων στο εμπειρικό τους υπόδειγμα εκφράστηκε σαν συνάρτηση της κατανάλωσης της προηγούμενης περιόδου, της τιμής των τσιγάρων και του ατομικού εισοδήματος. Όλες αυτές οι μεταβλητές μεταβάλλονται τόσο ως προς τον χρόνο όσο και ως προς τις πολιτείες.

Παρόλα αυτά υπάρχουν όμως αρκετοί παράγοντες οι οποίοι επηρεάζουν επίσης την κατανάλωση τσιγάρων οι οποίοι όμως δεν μεταβάλλονται ούτε ως προς τον χρόνο ούτε ως προς τις επιμέρους πολιτείες. Παραδείγματα τέτοιων μεταβλητών οι οποίες δεν μεταβάλλονται με το πέρασμα των ετών είναι το θρήσκευμα και το επίπεδο εκπαίδευσης του πληθυσμού σε κάθε πολιτεία. Από την άλλη πλευρά οι διαφημιστικές εκστρατείες των καπνοβιομηχανιών λαμβάνουν χώρα σε ομοσπονδιακό επίπεδο και επομένως δεν διαφοροποιούνται ως προς τις επιμέρους πολιτείες. Όμως και οι τρεις αυτές μεταβλητές αναμένουμε να επηρεάζουν a priori το επίπεδο κατανάλωσης τσιγάρων. Επιπλέον είναι πιθανόν να υπάρχουν κάποιες άλλες μεταβλητές οι οποίες να επηρεάζουν την κατανάλωση αλλά να μην είναι εύκολα μετρήσιμες ή προσβάσιμες. Ο αποκλεισμός όμως τέτοιων μεταβλητών από την συνάρτηση κατανάλωσης μπορεί να οδηγήσει σε σφάλματα κατά την οικονομετρική της εκτίμηση. Η χρήση panel data δίνει την δυνατότητα να

λάβουμε έμμεσα υπόψη μας κατά την οικονομετρική εκτίμηση του υποδείγματος όλες εκείνες τις μεταβλητές οι οποίες δεν μπορούν αντικειμενικά να συμπεριληφθούν στο εμπειρικό υπόδειγμα άμεσα.

Η χρήση panel data μας παρέχει περισσότερες πληροφορίες για τις οικονομικές μονάδες, περισσότερη διακύμανση στις μεταβλητές, λιγότερη συγγραμμικότητα μεταξύ των μεταβλητών, περισσότερους βαθμούς ελευθερίας και περισσότερη αποτελεσματικότητα στην οικονομετρική εκτίμηση. Τα οικονομετρικά υποδείγματα που βασίζονται αποκλειστικά σε στοιχεία χρονολογικών σειρών υποφέρουν σημαντικά από την ύπαρξη έντονης πολυσυγγραμμικότητας (multicollinearity). Στο παράδειγμα της κατανάλωσης τσιγάρων υπάρχει έντονη συγγραμμικότητα διαχρονικά μεταξύ της τιμής των τσιγάρων και του ατομικού εισοδήματος. Κάτι τέτοιο είναι όμως λιγότερο πιθανό όταν τα στατιστικά δεδομένα αφορούν και τις επιμέρους πολιτείες για την ίδια χρονική περίοδο καθώς η διαστρωματική διάσταση προσδίδει περισσότερη διακύμανση στις μεταβλητές καθώς και περισσότερες πληροφορίες αναφορικά με τις τιμές και το εισόδημα.

Η χρήση panel data μας δίνει την δυνατότητα να αναλύσουμε την διαδικασία και την δυναμική προσαρμογής των οικονομικών δεδομένων διαχρονικά διατηρώντας παράλληλα την ατομική ετερογένεια στα χαρακτηριστικά των οικονομικών μονάδων. Παραδείγματος χάρι για την μέτρηση της ανεργίας η χρήση απλών διαστρωματικών στοιχείων μπορεί να προσδιορίσει το επίπεδο ανεργίας σε κάθε δεδομένη χρονική στιγμή. Αντίθετα η χρήση panel data επιτρέπει την ανάλυση του επιπέδου ανεργίας για μία δεδομένη χρονική περίοδο και πώς αυτή συνδέεται με τα ιδιαίτερα χαρακτηριστικά των οικονομικών μονάδων.

Η ύπαρξη panel data επιτρέπει την εμπειρική εξειδίκευση περισσότερο πολύπλοκων θεωρητικών υποδειγμάτων για την ανάλυση της συμπεριφοράς των οικονομικών μονάδων η οποία είναι αδύνατη με την χρήση αποκλειστικά χρονολογικών σειρών ή διστρωματικών δεδομένων.

Εν ολίγοις, σε αντίθεση με τη χρησιμοποίηση μεμονωμένων διαστρωματικών δεδομένων ή δεδομένων χρονοσειρών, τα πάνελ δεδομένα μπορούν να εμπλουτίσουν την

εμπειρική ανάλυση. Παρ' όλα αυτά προβλήματα υπάρχουν στα μοντέλα παλινδρόμησης πάνελ δεδομένων, κάποια εκ των οποίων αναλύονται σε ενότητες που ακολουθούν. Τέλος, δεδομένου ότι τα στατιστικά δεδομένα συλλέγονται σε επίπεδο οικονομικής μονάδας μειώνεται σημαντικά το ενδεχόμενο σφάλματος κατά την διαδικασία άθροισης των δεδομένων σε εθνικό ή περιφερειακό επίπεδο.

### **3.2. ΓΕΝΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΟΣ ΣΕ PANEL DATA**

Στην οικονομετρία, η εκτίμηση μιας συνάρτησης παλινδρόμησης εξαρτάται από τις υποθέσεις που πραγματοποιούνται αναφορικά με τη σταθερά (intercept), τις κλίσεις (slope) της ευθείας καθώς και όρο σφάλματος (error term). Γενικά, οι περιπτώσεις που μπορεί να συναντήσει κανείς κατά την εμπειρική ανάλυση είναι αρκετές, ενδεικτικά αναφέρονται μερικές από αυτές στις ακόλουθες γραμμές:

- Η σταθερά και οι κλίσεις της ευθείας είναι σταθερές στο χώρο και στο χρόνο, ενώ ο όρος του σφάλματος διαφέρει με την πάροδο του χρόνου και μεταξύ των μονάδων – οντοτήτων
- Οι συντελεστές παλινδρόμησης είναι αμετάβλητοι, ενώ η σταθερά ποικίλει με τις οντότητες – μονάδες
- Οι συντελεστές παλινδρόμησης είναι αμετάβλητοι, ενώ η σταθερά ποικίλει με τις οντότητες – μονάδες και με την πάροδο του χρόνου
- Όλοι οι συντελεστές, τόσο η σταθερά όσο και οι κλίσεις της ευθείας ποικίλλουν μεταξύ των μονάδων οντοτήτων
- Η σταθερά από τη μια και οι συντελεστές παλινδρόμησης από την άλλη ποικίλλουν με την πάροδο του χρόνου και μεταξύ των μονάδων οντοτήτων.

Στο σημείο αυτό αξίζει να υπογραμμίσουμε ότι κάθε μια από τις προαναφερθείσες περιπτώσεις εμπεριέχει αυξημένη πολυπλοκότητα και ενσωματώνει ίσως περισσότερη αμεροληψία στην εκτίμηση των μοντέλων παλινδρόμησης πάνελ δεδομένων. Φυσικά η πολυπλοκότητα μπορεί να αυξηθεί ακόμη περισσότερο. Κάτι τέτοιο μπορεί να συμβεί αν

προστεθούν περισσότερες ανεξάρτητες μεταβλητές στο μοντέλο, εξαιτίας της δυνατότητας ύπαρξης συγγραμμικότητας μεταξύ αυτών.

Η αναλυτική ικανότητα του κλασσικού γραμμικού οικονομετρικού υποδείγματος βασίζεται στην υπόθεση ότι όλα τα δεδομένα ενός οικονομικού συστήματος είναι ελεγχόμενα με αποτέλεσμα όλες οι τυχαίες μεταβλητές να ακολουθούν μία ομαλή συνάρτηση κατανομής πιθανότητας. Δυστυχώς όμως τα οικονομικά δεδομένα δημιουργούνται από περισσότερο πολύπλοκες διαδικασίες που λαμβάνουν χώρα μέσα σε κάθε οικονομία. Η τυπική υπόθεση ότι κάθε οικονομική μεταβλητή  $y$  ακολουθεί μία παραμετρική συνάρτηση κατανομής πιθανότητας  $P(y|\theta)$  όπου  $\theta$  είναι το διάνυσμα των  $m$  παραμέτρων) η οποία είναι κοινή για όλες τις οικονομικές μονάδες μπορεί σε ορισμένες περιπτώσεις να μην είναι τόσο ρεαλιστική. Αγνοώντας όμως αυτή την ετερογένεια μεταξύ των παραμέτρων είναι πιθανό να οδηγηθούμε σε ανυπερείς ή ασήμαντες οικονομετρικές εκτιμήσεις.

Παραδείγματος χάρι ας θεωρήσουμε το παρακάτω γενικό γραμμικό υπόδειγμα:

$$y_{it} = \alpha_i^* + \beta_i x_{it} + v_{it}$$

Όπου,

$i = 1, 2, \dots, N$  είναι οι οικονομικές μονάδες και

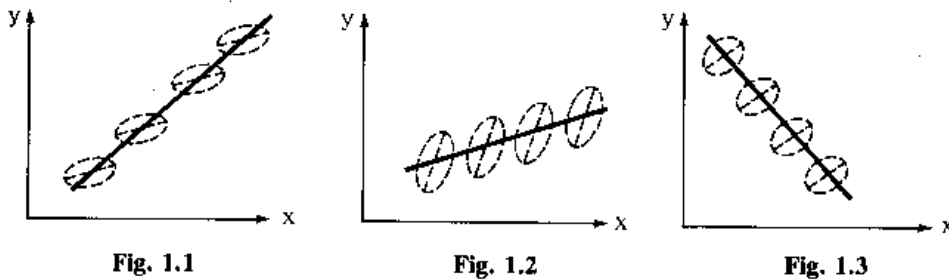
$t = 1, 2, \dots, T$  οι χρονικές περίοδοι για τις οποίες έχουμε παρατηρήσεις για αυτές,

$x_{it}$  είναι το διάνυσμα της ανεξάρτητης μεταβλητής και

$v_{it}$  είναι ο διαταρακτικός όρος του υποδείγματος ο οποίος ακολουθεί την κανονική κατανομή με μέσο μηδέν και σταθερή διακύμανση  $v_{it} \sim N(0, \sigma_v^2)$ .

Ας υποθέσουμε αρχικά ότι οι παράμετροι  $\alpha$  και  $\beta$  διαφέρουν μεταξύ των οικονομικών μονάδων αλλά παραμένουν σταθεροί κατά την διάρκεια της εξεταζόμενης περιόδου. Ακολουθώντας αυτή την υπόθεση μπορούν να προκύψουν από διαφορετικές κατανομές δειγματοληψίας. Αυτές οι διαφορετικές κατανομές δειγματοληψίας είναι πιθανό να οδηγήσουν σε ασυνεπείς εκτιμήσεις των παραμέτρων εάν αυτοί εκτιμηθούν με την απλή μέθοδο των ελαχίστων τετραγώνων χρησιμοποιώντας το σύνολο των  $NT$  παρατηρήσεων που έχουμε στη διάθεση μας. Ας διακρίνουμε τώρα δύο περιπτώσεις.

**Περίπτωση 1η:** ο σταθερός όρος της εξίσωσης παλινδρόμησης διαφέρει μεταξύ των οικονομικών μονάδων ( $\alpha_i^* \neq \alpha_j^*$ ) ενώ οι συντελεστές κλίσης παραμένουν σταθεροί ( $\beta_i = \beta_j$ )  $\forall i, j = 1, 2, \dots, N$ .



Σχήμα 13: Panel Data - Παλινδρόμηση<sup>11</sup>

Στα παραπάνω διαγράμματα οι ασυνεχείς κύκλοι παρουσιάζουν την περιοχή συγκέντρωσης των δεδομένων κάθε οικονομικής μονάδας (υποθέτουμε 4) κατά την διάρκεια της εξεταζόμενης χρονικής περιόδου. Οι ασυνεχείς γραμμές παρουσιάζουν τις εξισώσεις παλινδρόμησης για κάθε μία από τις οικονομικές μονάδες του παραδείγματος, ενώ η συνεχής γραμμή την ελαχίστων τετραγώνων εκτίμηση της εξίσωσης παλινδρόμησης εάν χρησιμοποιήσουμε όλες τις NT παρατηρήσεις (*pooled regression*). Όπως είναι εμφανές από τα παραπάνω διαγράμματα η χρήση όλων των παρατηρήσεων οδηγεί σε σημαντικά σφάλματα ως προς την εκτίμηση της εξίσωσης παλινδρόμησης η κατεύθυνση των οποίων δεν μπορεί να προβλεφθεί *a priori* με ακρίβεια.

**Περίπτωση 2η:** τόσο ο σταθερός όρος όσο και οι συντελεστές κλίσης διαφέρουν μεταξύ των οικονομικών μονάδων, δηλαδή ισχύει ( $\alpha_i^* \neq \alpha_j^*$ ) και ( $\beta_i \neq \beta_j$ )  $\forall i, j = 1, 2, \dots, N$ .

Τα παρακάτω διαγράμματα παρουσιάζουν τις εκτιμήσεις των ατομικών εξισώσεων παλινδρόμησης για δύο διαφορετικές περιπτώσεις κατανομής δειγματοληψίας. Και στις δύο περιπτώσεις η χρήση όλων των NT παρατηρήσεων για την εκτίμηση του υποδείγματος θα οδηγήσει σε ασυνεπή αποτελέσματα καθώς οι τιμές των εκτιμημένων

<sup>11</sup> <ftp://filer.soc.uoc.gr/>



παραμέτρων θα αποτελούν απλά τον μέσο όρο των ατομικών παραμέτρων. Και στις δύο περιπτώσεις η κλασική υπόθεση του Keynes για την ύπαρξη ενός αντιπροσωπευτικού καταναλωτή ή επιχείρησης δεν ευσταθεί.

Σχήμα 14: Panel Data - Σφάλματα<sup>12</sup>

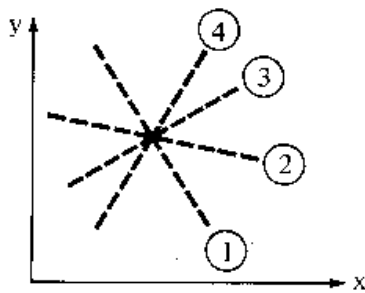


Fig. 1.4

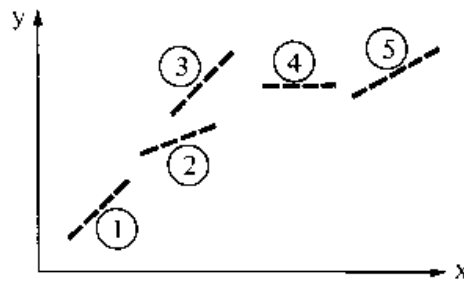


Fig. 1.5

Τα παραπάνω σφάλματα από την ύπαρξη ετερογενητικότητας στους παραμέτρους του υποδείγματος μπορούν να γίνουν πιο εμφανή εάν υποθέσουμε ότι αυτοί δεν παραμένουν σταθεροί ως προς τον χρόνο.

Το σφάλμα επιλεκτικότητας είναι επίσης πολύ συχνό στα οικονομετρικά υποδείγματα τα οποία βασίζονται στην χρήση είτε διαστρωματικών στοιχείων είτε *panel data*. Το σφάλμα επιλεκτικότητας αναφέρεται στις περιπτώσεις εκείνες όπου το χρησιμοποιούμενο δείγμα δεν επιλέγεται τυχαία από τον πληθυσμό. Ας υποθέσουμε ότι θέλουμε να μελετήσουμε την φορολογία εισοδήματος φυσικών προσώπων σε μία περιοχή της χώρας. Εάν μας ενδιαφέρουν μόνο τα χαμηλά εισοδήματα τότε το δείγμα το οποίο θα επιλέξουμε θα πρέπει να περιλαμβάνει φυσικά πρόσωπα των οποίων τα εισοδήματα δεν θα πρέπει να ξεπερνούν ένα ποσό έστω 1,5 φορά μεγαλύτερο από το όριο φτώχειας. Στην περίπτωση αυτή η χρήση του εισοδήματος σαν εξαρτημένη μεταβλητή σε μία εξίσωση παλινδρόμησης είναι πιθανό να οδηγήσει σε ασυνεπείς εκτιμήσεις των παραμέτρων (Hausman and Wise, 1977; Heckman, 1976; Hsiao, 1974).

<sup>12</sup> <ftp://filer.soc.uoc.gr/>

Ας θεωρήσουμε για λόγους απλούστευσης ότι χρησιμοποιούμε αποκλειστικά διαστρωματικά στοιχεία (η ανάλυση δεν διαφέρει εάν χρησιμοποιήσουμε *panel data*) ενώ το εισόδημα των νοικοκυριών εξαρτάται μόνο από το επίπεδο εκπαίδευσης. Τότε η εξίσωση παλινδρόμησης θα έχει την παρακάτω μορφή:

$$y_i = \alpha + \beta x_i + v_i$$

όπου

$y_i$  και  $x_i$  είναι το εισόδημα και το επίπεδο εκπαίδευσης των  $i = 1, 2, \dots, N$  νοικοκυριών και

$v_i$  ο διαταρακτικός όρος του υποδείγματος ο οποίος κατανέμεται κανονικά με μέσο μηδέν και σταθερή διακύμανση  $v_i \sim N(0, \sigma_v^2)$ .

Εάν όλα τα νοικοκυριά που συμμετέχουν στο δείγμα επιλέγονται με βάση το γεγονός ότι το εισόδημα τους είναι τουλάχιστον ίσο ή μικρότερο από 1,5 φορά το ελάχιστου επιπέδου διαβίωσης τότε το κριτήριο επιλογής με βάση την τελευταία σχέση μπορεί να γραφεί ως εξής:

$$y_i = \alpha + \beta x_i + v_i \leq L$$

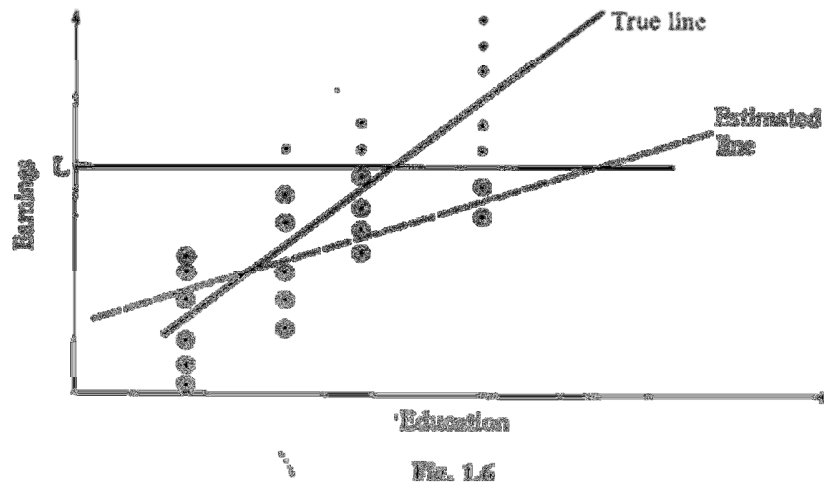
ή

$$y_i = \alpha + \beta x_i + v_i > L$$

Η των ελαχίστων τετραγώνων εκτίμηση της εξίσωσης παλινδρόμησης παρουσιάζεται στο παρακάτω διάγραμμα.

Η συνεχής γραμμή του διαγράμματος μας δίνει την αληθινή εξίσωση παλινδρόμησης μεταξύ εισοδήματος και εκπαίδευσης των νοικοκυριών. Εάν όμως επιλέξουμε από τον πληθυσμό δείγμα το οποίο πληροί το κριτήριο των παραπάνω εξισώσεων τότε η εκτιμημένη εξίσωση παλινδρόμησης έχει την μορφή της ασυνεχούς γραμμής στο διάγραμμα. Αποτελεί εμφανώς υποεκτίμηση της πραγματικής σχέσης

μεταξύ εισοδήματος και εκπαίδευσης. Με άλλα λόγια το σφάλμα επιλεκτικότητας δημιουργεί συσχέτιση μεταξύ των ανεξάρτητων μεταβλητών του υποδείγματος και του διαταρακτικού όρου γεγονός το οποίο οδηγεί σε υποεκτίμηση των παραμέτρων του υποδείγματος.



Σχήμα 15: Σφάλμα Επιλεκτικότητας<sup>13</sup>

### 3.3. ΜΟΝΤΕΛΟ ΜΗΔΕΝΙΚΩΝ ΕΠΙΔΡΑΣΕΩΝ

Το μοντέλο μηδενικών επιδράσεων αποτελεί μια απλουστευμένη διαδικασία, με βάση την οποία θεωρείται ότι δεν υπάρχει μη παρατηρούμενη ετερογένεια, δηλαδή ούτε διαστρωματικές ούτε χρονικές επιδράσεις ενώ οι συντελεστές που εκτιμώνται από το υπόδειγμα αυτό είναι σταθεροί τόσο ως προς το χρόνο όσο και ανά παρατηρούμενη μονάδα (διαστρωματικά). Επειδή είναι ιδιαίτερα σπάνιο να βρεθεί υπόδειγμα όπου τόσο οι διαστρωματικές όσο και οι χρονικές επιδράσεις να είναι στατιστικά ασήμαντες, το συγκεκριμένο μοντέλο δεν τυγχάνει ευρείας χρήσης. Μάλιστα, οφείλουμε να σημειώσουμε πως στην περίπτωση που υπάρχουν επιδράσεις και δεν ληφθούν υπόψη,

<sup>13</sup> <ftp://filer.soc.uoc.gr/>

τότε αφενός αλλοιώνεται η σχέση μεταξύ εξαρτημένης και ανεξάρτητων μεταβλητών, αφετέρου οι εκτιμητές ελαχίστων τετραγώνων θα είναι μεροληπτικοί και ασυνεπείς.

### 3.4. ΜΟΝΤΕΛΟ ΣΤΑΘΕΡΩΝ ΕΠΙΔΡΑΣΕΩΝ

Είναι ήδη γνωστό πως αν από το οικονομετρικό υπόδειγμα αγνοηθεί η μη παρατηρούμενη ετερογένεια και αν αυτή δεν είναι τυχαία αλλά συσχετίζεται με τις παρατηρούμενες ερευνητικές μεταβλητές, τότε οι συντελεστές των εκτιμητών θα είναι μεροληπτικοί και ασυνεπείς. Το πρόβλημα αυτό μπορεί να επιλυθεί μετασχηματίζοντας κατάλληλα το υπόδειγμα που χρησιμοποιούμε, ώστε να γίνει απαλοιφή των μη παρατηρούμενων επιδράσεων και αυτό επιτυγχάνεται με το γνωστό ως μετασχηματισμό σταθερών επιδράσεων (Fixed effects transformation).

Ο μετασχηματισμός αυτός έγκειται στη χρησιμοποίηση των αποκλίσεων των παρατηρήσεων από τον αντίστοιχο μέσο κάθε διαστρωματικής μονάδας. Από το μετασχηματισμένο πλέον υπόδειγμα θα γίνει απαλοιφή όχι μόνο της μη παρατηρούμενης ετερογένειας αλλά και του σταθερού όρου που πρεσβεύει το σύνολο των παρατηρούμενων επιδράσεων που παραμένουν σταθερές διαχρονικά (όπως για παράδειγμα φύλο, φυλή, κ.λπ.). η εκτίμηση του μετασχηματισμένου μοντέλου γίνεται με την κλασική μέθοδο των ελαχίστων τετραγώνων και οι εκτιμητές που προκύπτουν ονομάζονται σταθερών επιδράσεων εκτιμητές (fixed effects estimators).

Το Υπόδειγμα Απαρατήρητων ή Σταθερών Επιδράσεων (Fixed Effects Model) έχει την εξής μορφή:

$$Y_{it} = \beta_0 + \beta_1 X_{it,1} + \beta_2 X_{it,2} + \dots + \beta_k X_{it,k} + \alpha_i + u_{it}, \text{Cov}(\alpha_i, X_{it}) \neq 0$$

Το ζήτημα που προκύπτει επομένως, είναι η εκτίμηση του Υποδείγματος Απαρατήρητων Επιδράσεων, όταν η συνδιακύμανση των ανεξάρτητων μεταβλητών και των μη παρατηρούμενων παραγόντων που επηρεάζουν την εξαρτημένη μεταβλητή και οι οποίοι παραμένουν σταθεροί στο χρόνο είναι διάφορη του μηδενός. Η πρώτη σκέψη

είναι να ομαδοποιηθούν χρονικά οι παρατηρήσεις και να χρησιμοποιηθεί μέθοδος OLS. Όμως, λόγω του περιορισμού που προαναφέρθηκε δεν είναι εφικτό να προκύψουν συνεπείς εκτιμητές των παραμέτρων του υποδείγματος.

Τα κυριότερα Πλεονεκτήματα και οι αντίστοιχοι Περιορισμοί της Μεθόδου Σταθερών Επιδράσεων (Fixed Effects) συνοψίζονται στα ακόλουθα:

- Δυνατότητα για έλεγχο μη παρατηρήσιμων μεταβλητών που:
  - μεταβάλλονται δια - χρονικά και όχι δια - στρωματικά, ή/ και
  - μεταβάλλονται δια - στρωματικά και όχι δια - χρονικά
- Περισσότερες παρατηρήσεις παρέχουν περισσότερη πληροφόρηση
- Η εκτίμηση αφορά σε σχεδόν άμεσες επεκτάσεις της πολυμεταβλητής (πολλαπλής) παλινδρόμησης
- Η εκτίμηση σταθερών επιδράσεων γίνεται με τρεις τρόπους:
  - Μέθοδος «μεταβολών» για  $T = 2$
  - Μέθοδος « $n-1$  ερμηνευτικών μεταβλητών» για μικρό  $n$
  - Παλινδρόμηση «αποκλίσεων κάθε στρώματος από το χρονικό του μέσο»
- Παρόμοιες μέθοδοι εφαρμόζονται στην περίπτωση σταθερών επιδράσεων χρόνου όπως και στην περίπτωση σταθερών επιδράσεων χρόνου και στρώματος
- Η στατιστική επαγωγή γίνεται όπως και στην πολλαπλή παλινδρόμηση.

Αναφορικά με τους περιορισμούς στους οποίους υπόκειται το μοντέλο σταθερών επιδράσεων να σημειώσουμε πως οι κυριότεροι είναι οι ακόλουθοι:

- Ανάγκη για μεταβλητικότητα του  $X$  διαχρονικά «μέσα» στο στρώμα
- Οι επιδράσεις χρονικών υστερήσεων μπορεί να είναι σημαντικές
- Τα τυπικά σφάλματα μπορεί να είναι πολύ μικρά (ο διαταρακτικός όρος μπορεί να συσχετίζεται διαχρονικά).

Να σημειώσουμε στο σημείο αυτό πως κατά μια άλλη προσέγγιση, η μη παρατηρούμενη ετερογένεια θα μπορούσε να εκληφθεί όχι ως μη παρατηρούμενη μεταβλητή αλλά ως ξεχωριστή παράμετρος. Το υπόδειγμα στην περίπτωση αυτή μπορεί να εκτιμηθεί με την προσθήκη  $N - 1$  (για την αποφυγή τέλει συγγραμμικότητας)

ψευδομεταβλητών για τις διαστρωματικές μεταβλητές – μονάδες. Οι εκτιμητές που προκύπτουν καλούνται εκτιμητές ελαχίστων τετραγώνων (least squares estimators) ενώ με βάση τη βιβλιογραφία αποδεικνύεται ότι οι εκτιμητές σταθερών επιδράσεων και οι εκτιμητές ελαχίστων τετραγώνων είναι ίσοι μεταξύ τους.

Το πλεονέκτημα της χρήσης ψευδομεταβλητών είναι ότι με την εισαγωγή τους στο υπόδειγμα καταργείται η υπόθεση περί όμοιων χαρακτηριστικών μεταξύ των υπό εξέταση διαστρωματικών μονάδων, αφαιρούνται οι τυχαίες επιδράσεις και παράγονται εκτιμητές που βασίζονται μόνο στη χρονική διακύμανση των δεδομένων μέσα στην κάθε διαστρωματική μονάδα. Κατά αντίστοιχο τρόπο στο υπόδειγμα μπορούν να εισαχθούν και ψευδομεταβλητές για την εκτίμηση των χρονικών επιδράσεων (όπου στην περίπτωση αυτή γίνεται εισαγωγή  $T - 1$  ψευδομεταβλητών), και έτσι στο μοντέλο να υπάρχουν ταυτόχρονα και χρονικές ψευδομεταβλητές και ψευδομεταβλητές διαστρωματικών μονάδων.

Κριτήρια για το εάν στο υπόδειγμα θα εισαχθούν ψευδομεταβλητές της μιας ή της άλλης κατηγορίας ή και των δυο, αποτελούν η στατιστική σημαντικότητα των συντελεστών των μεταβλητών, η ερμηνευτική ικανότητα του υποδείγματος καθώς και το κριτήριο Durbin – Watson. Επίσης μπορεί να διεξαχθεί και ένας F έλεγχος στους συντελεστές των ψευδομεταβλητών όπου ως μηδενική υπόθεση εξετάζεται η περίπτωση οι συντελεστές όλων των ψευδομεταβλητών που έχουν εισαχθεί στο υπόδειγμα να είναι μηδέν, η εναλλακτική υπόθεση ισχύει αν τουλάχιστον ένας από αυτούς τους συντελεστές είναι διάφορος του μηδενός.

Το Μοντέλο Σταθερών Επιδράσεων (Fixed Effects Model) ή αλλιώς Μοντέλο Παλινδρόμηση Ελαχίστων Τετραγώνων με Ψευδομεταβλητές (Least Squares Dummy Variable Regression Model - LSDV), που αναλύθηκε στις παραγράφους που προηγήθηκαν είναι αρκετά απλοϊκό στη χρήση του, για αυτό και είναι τόσο διαδεδομένο. Από την άλλη πλευρά το μοντέλο αυτό παρουσιάζει και κάποια μειονεκτήματα, τα κυριότερα εκ των οποίων παρουσιάζονται ακολούθως:

- Εάν στο μοντέλο εισαχθούν αρκετές ψευδομεταβλητές τότε μπορεί να προκύψει το πρόβλημα των βαθμών ελευθερίας (δηλαδή η «χασούρα»

βαθμών ελευθερίας σχετικά με τις ψυδομεταβλητές, τις κλίσεις της ευθείας και την κοινή σταθερά)

- Ενσωματώνοντας πολλές μεταβλητές στο υπόδειγμα ελλοχεύει πάντα η πιθανότητα ύπαρξης πολυσυγγραμμικότητας η οποία μπορεί να καταστήσει την ακριβή εκτίμηση μιας ή περισσότερων παραμέτρων δύσκολη
- Εάν υποτεθεί ότι στο Μοντέλο Σταθερών Επιδράσεων συμπεριληφθούν μεταβλητές, όπως λ.χ. το φύλο, η εθνικότητα και το χρώμα (μεταβλητές σταθερές με την πάροδο του χρόνου), τότε η προσέγγιση LSDV ίσως να μην είναι σε θέση να προσδιορίσει την επίδραση τέτοιων αμετάβλητων στο χρόνο παραγόντων.

### **3.5. ΜΟΝΤΕΛΟ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ**

Σε συνέχεια των παραπάνω, στην παρούσα ενότητα εξετάζουμε το μοντέλο τυχαίων επιδράσεων. Η δεύτερη συχνά χρησιμοποιούμενη κατηγορία panel υποδειγμάτων είναι αυτά των τυχαίων επιδράσεων, όπου ο όρος *a<sub>it</sub>* είναι τυχαίος (random). Σε αντίθεση με το υπόδειγμα σταθερών επιδράσεων στο οποίο στόχος είναι να εξαλειφθεί η απαρατήρητη επίδραση, το υπόδειγμα τυχαίων επιδράσεων δεν προϋποθέτει κάτι τέτοιο αφού η σταθερή επίδραση δε συσχετίζεται με τις ερμηνευτικές μεταβλητές του υποδείγματος.

Σύμφωνα με αυτό το μοντέλο η μη παρατηρούμενη ετερογένεια εκλαμβάνεται ως τυχαία επίδραση οπότε και θεωρείται ότι δεν συσχετίζεται με τις παρατηρούμενες ερμηνευτικές μεταβλητές. Η βασική διαφορά αυτού σε σχέση με το μοντέλο σταθερών επιδράσεων είναι ότι ενώ σε αυτό η προσέγγιση προς το πραγματικό υπόδειγμα γίνεται μέσω της χρήσης ψευδομεταβλητών, πλέον με το μοντέλο τυχαίων επιδράσεων αυτή η προσέγγιση γίνεται με την ενσωμάτωση της τυχαίας επίδρασης στον διαταρακτικό όρο. Δηλαδή, σε αυτή την περίπτωση, η τυχαία επίδραση αποτελεί συστατικό μέρος του διαταρακτικού όρου, για το λόγο αυτό το υπόδειγμα το οποίο προκύπτει είναι επίσης γνωστό ως υπόδειγμα συστατικών σφάλματος (error components model).

Το Υπόδειγμα Τυχαίων Επιδράσεων (Random Effects Model) έχει την εξής μορφή:

$$Y_{it} = \beta_0 + \beta_1 X_{it,1} + \beta_2 X_{it,2} + \dots + \beta_k X_{it,k} + \alpha_i + u_{it}, \text{Cov}(\alpha_i, X_{it}) = 0$$

Από τα παραπάνω γίνεται εύκολα αντιληπτό πως για κάθε διαστρωματική μονάδα, η τιμή της συσχέτισης μεταξύ των διαταρακτικών όρων σε δυο διαφορετικές χρονικές στιγμές θα παραμείνει αμετάβλητη ανεξάρτητα από τη χρονική απόσταση των δυο αυτών χρονικών στιγμών. Επίσης η δομή της συσχέτισης παραμένει η ίδια για όλες τις διαστρωματικές μονάδες.

Επισημαίνεται ότι επειδή οι διακυμάνσεις των συστατικών μερών του διαταρακτικού όρου δεν είναι γνωστές, θα πρέπει να εκτιμηθούν. Η εφαρμογή της απλής μεθόδου των ελαχίστων τετραγώνων κρίνεται σε αυτή την περίπτωση ανεπαρκής εξαιτίας του μεγάλου αριθμού διαστρωματικών μονάδων που συνήθως περιλαμβάνουν οι εφαρμογές με δεδομένα Panel. Έτσι στην περίπτωση αυτή για την εκτίμηση των παραμέτρων χρησιμοποιείται η μέθοδος των γενικευμένων ελαχίστων τετραγώνων (generalized least squares method) μετά από κατάλληλο μετασχηματισμό του υποδείγματος για τη μείωση του μεγέθους της μήτρας διακυμάνσεων – συνδιακυμάνσεων του σύνθετου διαταρακτικού όρου.

Σε αυτό το σημείο αξίζει να αναφερθούμε σε δυο βασικά χαρακτηριστικά του συντελεστής συσχέτισης του συγκεκριμένου μοντέλου. Αρχικά, για κάθε εξεταζόμενη διαστρωματική μονάδα η τιμή του συντελεστή συσχέτισης μεταξύ των όρων σφάλματος σε δυο διαφορετικές χρονικές περιόδους, παραμένει σταθερή, ανεξάρτητα από το χρονικό διάστημα που λαμβάνεται υπόψη. Η παραπάνω διατύπωση έρχεται σε αντίθεση με την ανάλυση του αυτοπαλίνδρομου υποδείγματος πρώτης τάξης, AR(1), το οποίο έχει τονίσει ότι η συσχέτιση μεταξύ χρονικών περιόδων μειώνεται με την πάροδο του χρόνου. Επιπλέον, η τιμή της συσχέτισης παραμένει η ίδια για όλες τις διαστρωματικές μονάδες (δηλαδή, είναι σταθερή για όλες τις οντότητες).



### 3.6. ΣΥΓΚΡΙΣΗ ΜΟΝΤΕΛΩΝ ΣΤΑΘΕΡΩΝ ΚΑΙ ΤΥΧΑΙΩΝ ΕΠΙΔΡΑΣΕΩΝ

Αρκετές φορές, οι ερευνητές αντιμετωπίζουν το δίλλημα σχετικά με το ποιο από τα δυο μοντέλα που παρουσιάστηκαν στις ενότητες που προηγήθηκαν είναι το καλύτερο. Συγκεκριμένη απάντηση στη συγκεκριμένη ερώτηση – πρόκληση εξαρτάται από πληθώρα παραμέτρων, όπως η υπόθεση που γίνεται αναφορικά με την πιθανή συσχέτιση του συστατικού διαστρωματικού σφάλματος και των ανεξάρτητων μεταβλητών,  $X$ . Εάν υποθεθεί ότι το σφάλμα και οι μεταβλητές  $X$  συσχετίζονται, τότε το Μοντέλο Σταθερών Επιδράσεων ίσως δώσει περισσότερα αξιόπιστα αποτελέσματα.

Αναλυτικότερα, οι Judge et al. προβαίνουν σε ορισμένες χρήσιμες παρατηρήσεις αναφορικά με την επιλογή του ενός εκ των δυο μοντέλων. Οι επισημάνσεις αυτές είναι οι εξής:

- Εάν το  $T$  (συνολικός αριθμός χρονικών παρατηρήσεων) είναι μεγάλο και το  $n$  (αριθμός διαστρωματικών μονάδων) είναι μικρό, είναι πιθανόν να υπάρχει ελάχιστη διαφορά στις τιμές των παραμέτρων που εκτιμώνται με το Μοντέλο Σταθερών Επιδράσεων και το Μοντέλο Συστατικό Σφάλματος. Ως εκ τούτου, σε αυτή την περίπτωση, η επιλογή βασίζεται στην υπολογιστική ευκολία. Σε αυτό το πλαίσιο, το Μοντέλο Σταθερών Επιδράσεων ίσως είναι προτιμότερο.
- Όταν το  $n$  είναι μεγάλο και το  $T$  είναι μικρό, όλες οι εκτιμήσεις που λαμβάνονται από τις δυο μεθόδους μπορεί να διαφέρουν σημαντικά. Πιο συγκεκριμένα, εάν υποθεθεί ότι οι διαστρωματικές μονάδες στο δείγμα δεν είναι τυχαία σχεδιασμένες από έναν μεγαλύτερο πληθυσμό, τότε το Μοντέλο Σταθερών Επιδράσεων είναι το καλύτερο δυνατό. Ωστόσο, εάν οι οντότητες στο δείγμα έχουν παρθεί τυχαία, τότε το Μοντέλο Τυχαίων Επιδράσεων είναι κατάλληλο.

- Εάν το διαστρωματικό τυχαίο συστατικό και μια ή περισσότερες ανεξάρτητες μεταβλητές,  $X$  συσχετίζονται, τότε οι εκτιμητές του Μοντέλου Συστατικών Σφάλματος είναι μεροληπτικοί, ενώ εκείνοι που λαμβάνονται από το Μοντέλο Σταθερών Επιδράσεων είναι αμερόληπτοι.
- Εάν το  $n$  είναι μεγάλο, το  $T$  είναι μικρό και εάν οι υποθέσεις στις οποίες βασίζεται το Μοντέλο Τυχαίων Επιδράσεων ισχύουν, τότε οι εκτιμητές του συγκεκριμένου μοντέλου είναι περισσότερο αποτελεσματικοί από ότι εκείνοι του Μοντέλου Σταθερών Επιδράσεων.

Είναι γεγονός πως προκειμένου οι ερευνητές να βοηθηθούν στην επιλογή του καταλληλότερου μοντέλου ο Hausman ανέπτυξε έναν στατιστικό έλεγχο, ο οποίος περιγράφεται εκτενώς στην ενότητα που ακολουθεί.

### 3.7. ΕΛΕΓΧΟΣ HAUSMAN

Σε μια εμπειρική μελέτη βασικό ζήτημα που τίθεται είναι το ποιο μοντέλο είναι καταλληλότερο να εφαρμοστεί. Ιδιαίτερη σημασία σε αυτό το θέμα έχει η ύπαρξη ή όχι συσχέτισης μεταξύ του όρου σφάλματος των διαστρωματικών μονάδων και των συντελεστών των ανεξάρτητων μεταβλητών. Εάν δεν υπάρχει συσχέτιση, τότε καταλληλότερο προς εφαρμογή είναι το μοντέλο των τυχαίων επιδράσεων.

Αντίθετα στην περίπτωση που εξακριβωθεί η ύπαρξη συσχέτισης, τότε το μοντέλο που χρησιμοποιείται είναι αυτό των σταθερών επιδράσεων. Πάνω σε αυτό το ζήτημα, ο Hausman (1978) πρότεινε την εφαρμογή ενός τεστ – ελέγχου βάσει του οποίου η επιλογή του μοντέλου θα μπορούσε πλέον να πραγματοποιηθεί με πιο αξιόπιστα κριτήρια.

Στο τεστ του Hausman χρησιμοποιείται η κατανομή  $X^2$  η ελεγχοσυνάρτηση χρησιμοποιείται για να αποδεχθεί ή να απορρίψει τη μηδενική υπόθεση, όπου ως μηδενική υπόθεση θεωρούμε τη μη ύπαρξη σημαντικών διαφορών μεταξύ των εκτιμώμενων συντελεστών των δυο μοντέλων. Εάν η μηδενική υπόθεση απορριφθεί, τότε πιο κατάλληλο να εφαρμοστεί κρίνεται το μοντέλο των σταθερών επιδράσεων. Αυτό που

κατά κύριο λόγο εξετάζεται με τον έλεγχο – τεστ του Hausman είναι η διαφορλα των εκτιμητριών των παραμέτρων  $\beta$  μεταξύ των δυο μοντέλων.

Αναλυτικότερα οι υποθέσεις που γίνονται κατά τον έλεγχο Hausman είναι οι ακόλουθες:

***H<sub>0</sub>**: ο συντελεστής  $\beta$  του μοντέλου των σταθερών επιδράσεων είναι συνεπής αλλά αναποτελεσματικός ενώ ο συντελεστής  $\beta$  του μοντέλου τυχαίων επιδράσεων είναι συνεπής και αποτελεσματικός*

***H<sub>1</sub>**: ο συντελεστής  $\beta$  του μοντέλου σταθερών επιδράσεων είναι συνεπής και ο συντελεστής  $\beta$  του μοντέλου τυχαίων επιδράσεων είναι μη συνεπής.*

Ολοκληρώνοντας την παρούσα ενότητα αλλά και το κεφάλαιο ανάλυσης των δεδομένων πάνελ, αξίζει να αναφέρουμε πως αν και ο έλεγχος Hausman χρησιμοποιείται κατά κόρον στη διεθνή βιβλιογραφία μπορούν να βρεθούν αρκετές ακόμη μελέτες όπου γίνεται χρήση εναλλακτικών μεθόδων επιλογής του καταλληλότερου μοντέλου με βάση ορισμένα εμπειρικά κριτήρια. Πάντως αυτό που επισημαίνεται από πληθώρα ερευνητών είναι ότι επειδή στην πράξη είναι δύσκολο να διαπιστωθεί η μη ύπαρξη συσχέτισης μεταξύ των ερμηνευτικών μεταβλητών και της ετερογένειας των διαστρωματικών μονάδων, η χρήση του μοντέλου σταθερών επιδράσεων αποτελεί στην πραγματικότητα την πιο κατάλληλη επιλογή.

## **ΚΕΦΑΛΑΙΟ 4 «ΕΦΑΡΜΟΓΕΣ ΑΝΑΛΥΣΗΣ ΔΕΔΟΜΕΝΩΝ ΠΑΝΕΛ – E-VIEWS»**

Το παρόν και τελευταίο κεφάλαιο της εν λόγω εργασίας, αφορά ορισμένα παραδείγματα εφαρμογής όπου απαιτείται η ανάλυση δεδομένων πάνελ. Το οικονομετρικό πακέτο EVIEWS και η χρησιμότητα αυτού ολοκληρώνουν το παρόν κεφάλαιο.

### **4.1. ΕΦΑΡΜΟΓΗ ΣΤΗ ΝΑΥΤΙΛΙΑ - ΠΑΡΑΔΕΙΓΜΑ**

Η ναυτιλία είναι ένας τομέας που αναγνωρίζει ιδιαίτερη ανάπτυξη και η ανάλυση αυτού είναι ιδιαίτερα σημαντική. Οι οικονομικές διακυμάνσεις που διέπουν τον κλάδο είναι αυτές που χρήζουν οικονομετρικής ανάλυσης και συγκεκριμένα ανάλυσης δεδομένων πάνελ.

Σαν οικονομικές διακυμάνσεις εννοούμε κυρίως τις περιοδικές ή ρυθμικές μεταβολές στην οικονομική δραστηριότητα ( αλλά και στις τιμές δηλ. στο γενικό επίπεδο των τιμών) ή και άλλων οικονομικών μεταβλητών. Η συνθήκη της περιοδικότητας δεν είναι αναγκαία. Ο όρος οικονομική δραστηριότητα περιλαμβάνει τη συνολική παραγωγή αγαθών και υπηρεσιών ή άλλως το πραγματικό εισόδημα της Οικονομίας το οποίο επηρεάζει τη συνολική απασχόληση.

Αρχικά υπήρξε η άποψη μεταξύ των επιστημόνων ότι οι οικονομικές διακυμάνσεις γνωστές αργότερα και ως κυκλικές διακυμάνσεις ήταν ακανόνιστες και τυχαίες όπως τα φυσικά φαινόμενα και ότι δεν υπόκειντο σε κανενός είδους πρόβλεψη. Αργότερα επικράτησε η αντίληψη ότι στις οικονομικές διακυμάνσεις υπάρχει περιοδικότητα ή ρυθμικότητα διότι παρατηρήθηκε το φαινόμενο της επανεμφάνισης των οικονομικών διακυμάνσεων κατά χρονικά διαστήματα λίγο ή πολύ σταθερά.

Είναι κοινώς αποδεκτό ότι η οικονομική δράση των οικονομιών και των ατόμων μέσα σε αυτές παρουσιάζει κυκλικές εναλλασσόμενες φάσεις ανόδου και καθόδου. Στο σημείο αυτό εμπλουτίστηκε η έρευνα περαιτέρω με τη μελέτη της λεγόμενης μακροαίωνης τάση (secular trend) και των εποχικών διακυμάνσεων. Σχετικά με την μακροαίωνη τάση αυτή δεν παρουσιάζει κυκλικές διακυμάνσεις αλλά οι εναλλασσόμενες φάσεις της ανόδου και καθόδου είναι σταθερές και συνεχείς.

Στις οικονομικές διακυμάνσεις περιλαμβάνονται τόσο οι διαρθρωτικές αλλαγές στην οικονομία όσο και σε τυχαία γεγονότα οφειλόμενες διακυμάνσεις (όπως οι μεταναστεύσεις πληθυσμών, τα επίπεδα γεωγραφικής παραγωγής, η κρίση ή κάμψη του Χρηματιστηρίου και τα συναφή με αυτά).

Στην περίπτωση της ναυτιλίας εξετάζουμε τους ναυτιλιακούς κύκλους και γενικά η κυκλική φύση της αγοράς χύδην ξηρών φορτίων δεν μπορεί να επισημανθεί αρκετά με κορυφώσεις και πτώσεις των τιμών των ναύλων να συμβαίνουν με αλάνθαστη συχνότητα στην πάροδο των τελευταίων 20 χρόνων. Εκτός από συγκεκριμένες εξαιρέσεις αυτή η κυκλική συμπεριφορά είναι αυτή που τελικά θα καθορίσει τα γεγονότα στο μέλλον. Εν συντομία επομένως η ιστορία της αγοράς χύδην ξηρών φορτίων επιβεβαιώνει το ρητό ότι αυτοί που αποτυγχάνουν να μάθουν τα μαθήματα της ιστορίας είναι αναγκασμένοι να τα επαναλάβουν.

- Η έναρξη της ανόδου παραδοσιακά προκαλείται από αυξημένα επίπεδα ζήτησης σε συνδυασμό με στασιμότητα ή μείωση των παραγγελιών νέων πλοίων και αύξηση των πλοίων που πηγαίνουν για scrap
- Η κορύφωση όπου στο σημείο αυτό ο μεγάλος όγκος για παραγγελίες νέων πλοίων αρχίζει να προκαλεί φόβο για υπερπροσφορά χωρητικότητας
- Η πτώση της αγοράς πραγματοποιείται όταν οι παραγγελίες για νέα πλοία σταματούν οι πωλήσεις πλοίων για scrap αυξάνονται και η διαδικασία διόρθωσης μπαίνει σε λειτουργία

Οι ναυτιλιακοί κύκλοι αποτελούν αναπόσπαστο κομμάτι της ναυτιλιακής βιομηχανίας γενικά και της αγοράς ναύλων ειδικότερα. Αν αναλογιστεί κανείς τα τεράστια ποσά που εμπλέκονται στο χώρο της ναυτιλίας δεν είναι δύσκολο να αντιληφθεί

τη σπουδαιότητα των ναυτιλιακών κύκλων στον καθορισμό των κρίσιμων επιχειρηματικών αποφάσεων. Το πρόβλημα του κόστους της θαλάσσιας μεταφοράς δεν αφορά μόνο τους πλοιοκτήτες. Οι φορτωτές / ναυλωτές αντιμετωπίζουν παρόμοιους προβληματισμούς.

Κατά τη διάρκεια οποιουδήποτε έτους και γενικά ανεξάρτητα από το θεμελιώδη ναυτιλιακό κύκλο η αγορά αναμένει ορισμένες εποχιακές μεταβολές να συμβούν. Μιλώντας γενικά οι παραδοσιακές φάσεις αφορούν ανάκαμψη την άνοιξη, καλοκαιρινή νηνεμία και ανοδική πορεία το χειμώνα. Η αυξημένη δραστηριότητα την άνοιξη οφείλεται κυρίως στα δημητριακά που μεταφέρονται κυρίως από Βραζιλία και Αργεντινή και σε ένα βαθμό και από την Αυστραλία. Συνήθως οι προσδοκίες για αυτή τη δραστηριότητα ( και τους ναύλους ) κορυφώνεται τον Απρίλιο ή τον Μάιο.

Η θερινή περίοδος τείνει να δει την αγορά να μπαίνει σε ένα βαθμό σε λήθαργο. Εν μέρει αυτό οφείλεται στις περιόδους διακοπών στο βόρειο ημισφαίριο, χρόνο που εκμεταλλεύονται μερικοί για να κλείσουν συμφωνίες ή να χρησιμοποιήσουν την περίοδο αυτή για να αναλάβουν την απαραίτητη εργασία της συντήρησης. Η αγορά σιταριού είναι επίσης σε εκκρεμότητα με την επιχείρηση που τείνει να στραφεί στην τελευταία της συγκομιδή της παλαιάς εποχής. Εκτός των εξαιρετικών καιρικών συνθηκών ( που σε μερικές περιοχές έχουν ως αποτέλεσμα την υψηλή κατανάλωση ενέργειας εξαιτίας της χρήσης κλιματιστικών ) αυτή τείνει επίσης να είναι η περίοδος του έτους όπου μειώνεται η κατανάλωση ενέργειας και ως εκ τούτου τα αποθέματα άντλησης βρίσκονται σε χαμηλά επίπεδα.

Τους φθινοπωρινούς μήνες έχουμε τις συγκομιδές σιταριού στο βόρειο ημισφαίριο μαζί με μια συγκέντρωση στη ζήτηση από ενέργεια. Ιστορικά υπήρχε η τάση να δίνεται σημασία στο επίπεδο κλεισίματος που είναι ικανό να καλύψει τις τελευταίες μέρες εξόδου από το σύστημα των Μεγάλων Λιμνών.

Σε ένα μικρό – οικονομικό επίπεδο υπάρχουν επίσης αναμενόμενες κινήσεις της αγοράς που επιτρέπουν στους πλοιοκτήτες και στους ναυλωτές να αποφύγουν να έχουν τονάζ ή επιχειρηματικές δραστηριότητες ανοιχτές κατά τη διάρκεια των γιορτών. Βασιζόμενος στις θεμελιώδεις κατευθύνσεις της εκάστοτε περιόδου κάποιος μπορεί να

είναι διατεθειμένος να μειώσει τις τιμές για να διατηρήσει σταθερή τη δουλειά. Τα Χριστούγεννα και την Πρωτοχρονιά είναι ένα τυπικό παράδειγμα.

Οι ναύλοι στην ναυτιλία αντικατοπτρίζουν την ισορροπία μεταξύ ζήτησης και προσφοράς. Η ζήτηση για μεταφορικές υπηρεσίες είναι παράγωγο ζήτησης όπου εξαρτάται :

- Τα οικονομικά των αγαθών που μεταφέρονται δια θαλάσσης, αυτό είναι το επίπεδο παραγωγής και της κατανάλωσης για το αγαθό που μεταφέρεται
- Τις παγκόσμιες οικονομικές συνθήκες και την παγκόσμια δραστηριότητα
- Τις σχετιζόμενες μικροοικονομικές μεταβλητές των κύριων οικονομιών

Αυτές οι μικροοικονομικές μεταβλητές έχουν τυχαίες μεταβολές και εποχιακά στοιχεία στις περισσότερες περιπτώσεις. Οι τιμές της αγοράς σε αρκετά αγαθά φαίνεται να εμφανίζουν εποχιακά χαρακτηριστικά, για παράδειγμα υπάρχουν οι εποχιακές μεταβολές στο εμπόριο σιταριών. Για αυτό το λόγο είναι πιθανό ότι αυτές οι εποχιακές μεταβολές να επηρεάσουν τις τιμές των ναύλων.

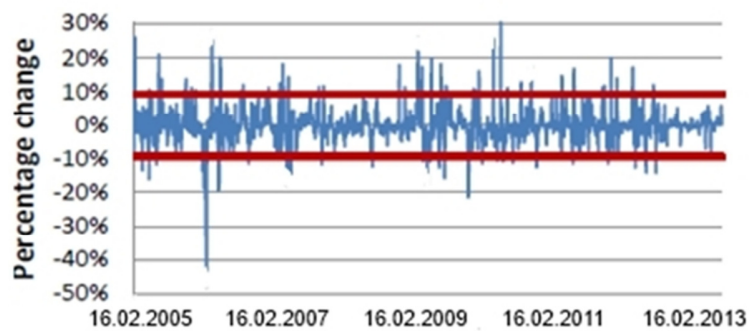
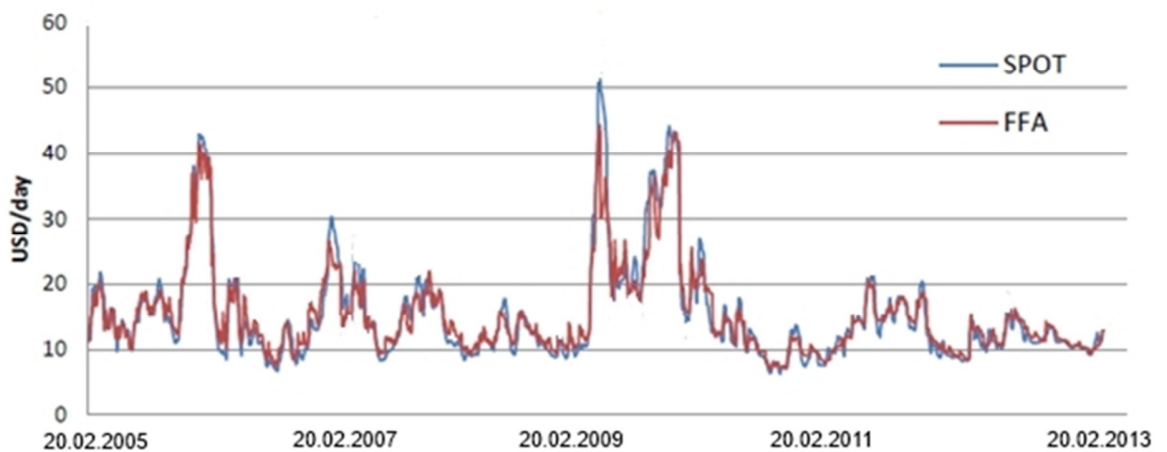
Παρακάτω παρουσιάζονται ορισμένα δεδομένα σχετικά με τη μεταβλητότητα των ναύλων, πραγματοποιείται εφαρμογή αυτών σε δεδομένα πάνελ. Για την εμπειρική ανάλυση που ακολουθεί τα δεδομένα αντλούνται από τη βάση δεδομένων Clarkson, χρονικά εκτείνονται από 14/02/2005 μέχρι 25/02/2013. Επιπλέον βάσεις δεδομένων που χρησιμοποιήθηκαν είναι η Baltic Maritime Exchange που παρέχει τα στοιχεία για τις TC2, TD3, TD5 και TD7, καθώς και η Platts που παρέχει τα δεδομένα για TC5. Τα στοιχεία τα οποία αναλύονται αφορούν FFA και SPOT αποδόσεις.

Ως περίοδος εξέτασης επιλέχθηκε η τελευταία οκταετία έτσι ώστε να περιλαμβάνονται στην ανάλυσή μας δεδομένα πριν και μετά την οικονομική κρίση, η οποία εκδηλώθηκε σε παγκόσμιο επίπεδο τον Οκτώβριο του 2008. Οι τιμές των FFA αναφέρονται σε Worldscale σημεία μέχρι και το τέλος του 2012, από το 2013 τα ημερήσια στοιχεία αναφέρονται στο USD / mt την ημέρα αντί των Worldscale σημείων.

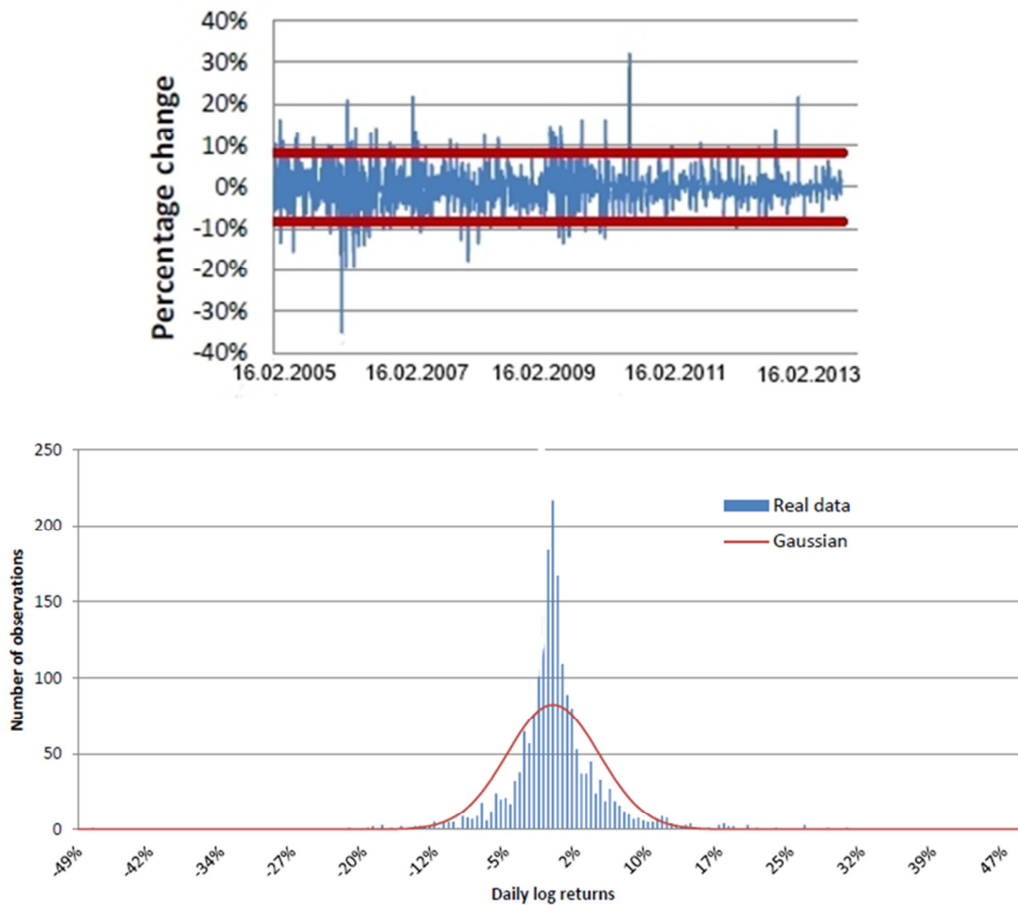
Τα τελευταία χρόνια παρατηρείται μια στοιχειώδης σταθεροποίηση στις τιμές των FFA σε επίπεδα χαμηλότερα από εκείνα πριν από την κρίση. Ένα επιπλέον σημαντικό συμπέρασμα που προκύπτει από τις παραπάνω εικόνες είναι πως οι αποδόσεις των FFA

όπως τα περισσότερα χρονολογικά στοιχεία παρουσιάζουν τα χαρακτηριστικά της κυκλικότητας και της εποχικότητας, γεγονός που ισχύει τόσο πριν όσο και μετά το 2008, όταν και σημειώθηκε η παρούσα χρηματοοικονομική κρίση.

Οι διαδρομές που αναλύονται είναι οι: TD3, TD5, επιλέχθηκαν οι συγκεκριμένες διαδρομές προκειμένου η εμπειρική ανάλυση να συνάδει με την θεωρητική προσέγγιση που αφορά τα χύδην φορτία. Πραγματοποιείται αφενός περιγραφική ανάλυση αυτών και αφετέρου ανάλυση κινδύνου. Τα διαγράμματα που αφορούν τη διαδρομή TD3 παρατίθενται ακολούθως:







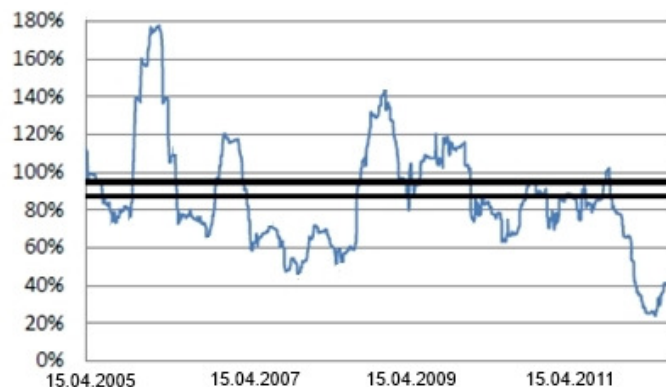
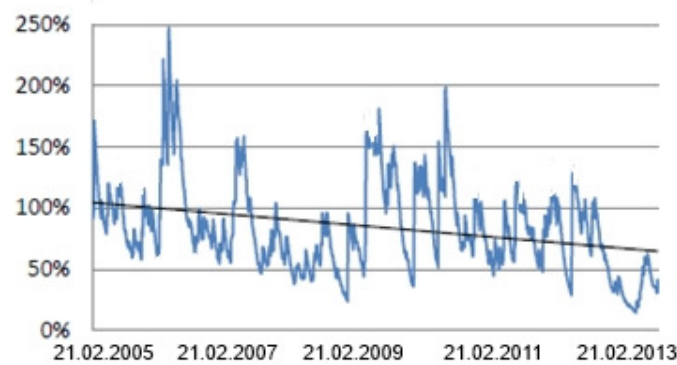
Σχήμα 16: Περιγραφική Ανάλυση Κινδύνου<sup>14</sup>

Στην παρούσα ενότητα θα παρουσιάσουμε τα αποτελέσματα που αφορούν την μεταβλητότητα των αποδόσεων στη διαδρομή TD3 τόσο για τα δεδομένα των SPOT όσο και για τις αποδόσεις των FFA.

Για τη μελέτη της μεταβλητότητας χρησιμοποιούμε την μέθοδο EWMA η οποία υποδεικνύει πως πρόκειται για στοχαστική μεταβλητή. Κατά τη διάρκεια της δειγματοληψίας υπολογίστηκε πως η μεταβλητότητα των SPOT παρουσιάζει έντονη μεταβλητότητα από 250% μέχρι κάτω από 50%. γεγονός σύμφωνο με την θεωρία που απαιτεί μεγάλη μεταβλητότητα για τα πλοία τύπου VLCC. Προσεγγίζοντας την

<sup>14</sup> [www.clarkson.com](http://www.clarkson.com)

μεταβλητότητα με την διαδικασία του Rolling Window διαπιστώνουμε πως παρουσιάζει τις ίδιες τάσεις με αυτές της μεθόδου EWMA. Να σημειωθεί στο σημείο αυτό πως υπολογίζοντας την μεταβλητότητα με αυτό τον τρόπο οι κορυφές τείνουν να είναι λιγότερο διακριτές σε σύγκριση με τη μέθοδο EWMA, αυτό συμβαίνει κυρίως επειδή οι ακραίες αποδόσεις έχουν την ίδια βαρύτητα με τις προηγούμενα 89 παρατηρήσεις. Υπολογίζοντας ένα 99% διάστημα εμπιστοσύνης στην περίπτωση της Rolling Window προσέγγισης διαπιστώνουμε πως η μεταβλητότητα υπερβαίνει τα χρονικά όρια γεγονός που υποδεικνύει πως η μεταβολή στην μεταβλητότητα δεν οφείλεται σε σφάλμα δειγματοληψίας. Και σε αυτή την περίπτωση επιβεβαιώνεται αυτό το οποίο υπογραμμίσαμε από την αρχή, ότι η μεταβλητότητα είναι στοχαστική μεταβλητή. Όλα τα παραπάνω γραφικά απεικονίζονται ως εξής:

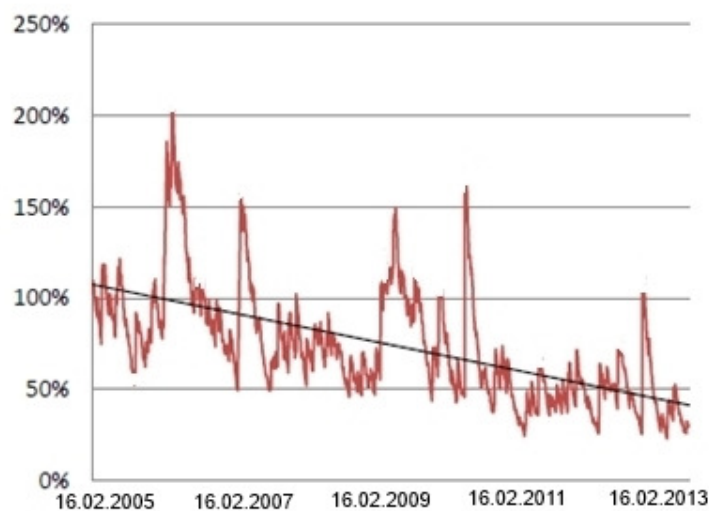


Σχήμα 17: Μεταβλητότητα Χρονοσειράς<sup>15</sup>

<sup>15</sup> [www.euro2day.gr](http://www.euro2day.gr)

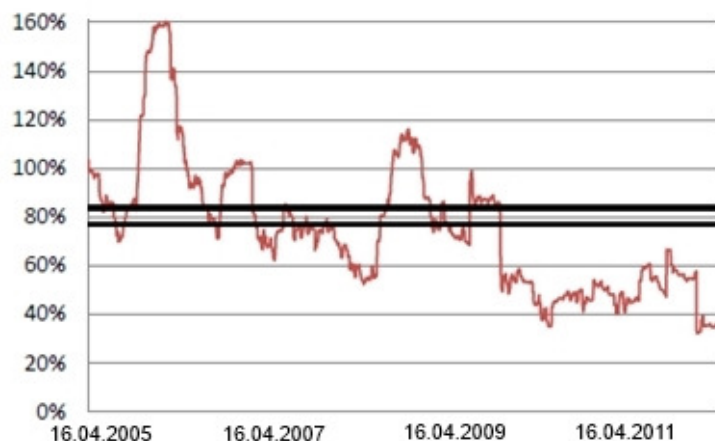
Αξίζει στο σημείο αυτό να σχολιάσουμε πως παρατηρείται μια πτωτική τάση της μεταβλητότητας στην υποεξέταση περίοδο γεγονός που πιθανόν να αιτιολογείται από την πλεονάζουσα προσφορά στην αγορά των tankers. Τέλος πολλές από τις κορυφές στη μεταβλητότητα πιθανότατα να οφείλονται σε εποχικότητα που σχετίζεται με τις διακοπές των Χριστουγέννων.

Σε συνέχεια των παραπάνω πραγματοποιούμε και ανάλυση μεταβλητότητας των FFA από όπου διαπιστώνεται πως σε σύγκριση με την μεταβλητότητα των SPOT παρουσιάζουν τις ίδιες τιμές ακόμη και στην εποχικότητα. Στην πραγματικότητα φαίνεται πως η μεταβλητότητα των FFA είναι οριακά μικρότερη τα τελευταία χρόνια της ανάλυσης με κάποιες έντονες διακυμάνσεις το 2012. Και σε αυτή την περίπτωση η μεταβλητότητα δεν οφείλεται σε σφάλμα δειγματοληψίας.



Σχήμα 18: Μεταβλητότητα FFA<sup>16</sup>

<sup>16</sup> [digilib.lib.unipi.gr/dspace/bitstream/unipi/5188](http://digilib.lib.unipi.gr/dspace/bitstream/unipi/5188)



Σχήμα 19: Μεταβλητότητα FFA ( $\alpha$ )<sup>17</sup>

Προκειμένου να ολοκληρωθεί η ανάλυση της μεταβλητότητας της διαδρομής TD3 παρουσιάζεται στον ακόλουθο πίνακα η μεταβλητότητα της μεταβλητότητας τόσο των SPOT όσο και των FFA, σε ετήσια βάση:

<b>Volatility of the volatility</b>				
	<b>SPOT daily</b>	<b>SPOT annually</b>	<b>FFA daily</b>	<b>FFA annually</b>
<b>2005 - 2013</b>	1.55%	29.68%	1.42%	27.28%
<b>2005 - 2006</b>	1.80%	35.41%	1.47%	28.95%
<b>2007 - 2008</b>	1.14%	21.32%	0.68%	13.98%
<b>2009 - 2010</b>	1.06%	20.97%	1.16%	21.78%
<b>2010 - 2011</b>	1.23%	23.32%	0.42%	7.37%
<b>2012 - 2013</b>	1.40%	25.20%	0.47%	9.87%

<sup>17</sup> [digilib.lib.unipi.gr/dspace/bitstream/unipi/5188/1/](http://digilib.lib.unipi.gr/dspace/bitstream/unipi/5188/1/)

## **4.2. ΕΦΑΡΜΟΓΗ ΣΤΟ ΧΡΗΜΑΤΙΣΤΗΡΙΟ (ΕΚΤΙΜΗΣΗ ΜΕΤΟΧΩΝ) - ΠΑΡΑΔΕΙΓΜΑ**

Τεχνική ανάλυση μιας μετοχής ή ενός Δείκτη, ή μιας ολόκληρης οργανωμένης αγοράς, είναι η μελέτη της κίνησης των τιμών, όπως αυτή απεικονίζεται σε ένα διάγραμμα τιμής- χρόνου, με στόχο τον χαρακτηρισμό της συμπεριφοράς της και την πρόβλεψη της μελλοντικής της κίνησης. Στηρίζεται στα γραφήματα της Αναλυτικής Γεωμετρίας και της Στατιστικής, τα οποία όμως έχουν εκλαϊκευτεί, για να μπορεί να τα καταλαβαίνει ο μέσος επενδυτής. Η Τεχνική Ανάλυση δεν ασχολείται με τα θεμελιώδη στοιχεία μιας εταιρείας ή μιας αγοράς, όπως είναι οι ισολογισμοί, ο λόγος P/E, ή τα κέρδη, αλλά μόνο με τα γραφήματα των μετοχών (ή των δεικτών).

Βασική υπόθεση της τεχνικής ανάλυσης είναι πως ότι είναι γνωστό για μια εταιρεία είναι ήδη ενσωματωμένο στην τιμή της. Για την τεχνική ανάλυση δεν έχει σημασία τι προκάλεσε την άνοδο ή κάθοδο της τιμής μιας μετοχής, αλλά πόσο θα διαρκέσει αυτή η μεταβολή, είτε προς τα πάνω είτε προς τα κάτω. Επίσης βασική παραδοχή της τεχνικής ανάλυσης αποτελεί το γεγονός ότι η τιμή μιας μετοχής απεικονίζει τα οικονομικά δεδομένα, τις πληροφορίες και τα νέα της και επηρεάζεται σημαντικά από τα ανθρώπινα συναισθήματα και την ψυχολογία που επικρατεί κάθε περίοδο.

Η Τεχνική Ανάλυση μελετάει τις μεταβολές τιμών από το παρελθόν και προσπαθεί να εντοπίσει έγκαιρα τις μεταβολές που θα γίνουν στο μέλλον και κυρίως το σημείο αντιστροφής μιας κίνησης τιμών.

Θεωρώντας δεδομένο ότι δεν γνωρίζουμε τι θα συμβεί στο μέλλον, με την Τεχνική Ανάλυση αναζητούμε πιθανές συμπεριφορές στο παρελθόν που μπορεί να μας βοηθήσουν ώστε να εφαρμόσουμε μια επενδυτική στρατηγική, βασιζόμενοι στο σκεπτικό ότι οι συμπεριφορές τείνουν να επαναλαμβάνονται με παρόμοιο τρόπο.

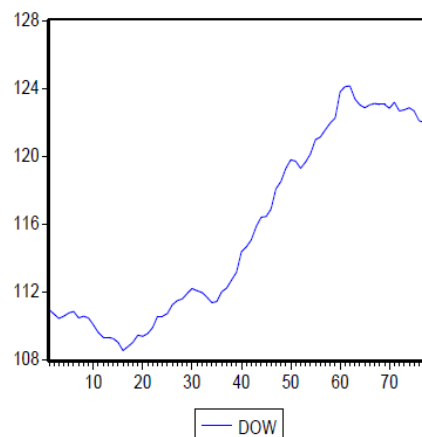
Πρέπει να γίνει κατανοητό σε όποιον δραστηριοποιείται στην αγορά, με εργαλείο την Τεχνική Ανάλυση ότι οι αγορές δεν κινούνται από τα γεγονότα, αλλά από τις προσδοκίες των ανθρώπων. Οι δε τιμές των μετοχών στο ταμπλό του χρηματιστηρίου δεν

απεικονίζουν το "σήμερα" αλλά το "αύριο" με ορίζοντα περίπου 3-6 μηνών ή και περισσότερο.

Κάθε αγορά μετοχών, παραγώγων, αξιών ή/και καταναλωτικών ειδών γεννά, με τη συμπεριφορά της, στους εμπλεκόμενους τα απολύτως απαραίτητα ανθρώπινα συναισθήματα της απληστίας, του φόβου, της ανάγκης μιμητισμού των άλλων και ένταξης σε μεγαλύτερα σύνολα, τα οποία αποτελούν τις αιτίες αποτυχημένης δραστηριοποίησης στην αγορά.

Οι Τεχνική Ανάλυση, θεωρεί ότι τα ανθρώπινα συναισθήματα απεικονίζονται στους σχηματισμούς των τιμών, και προσπαθεί να απομονώσει τις ψυχολογικές επιδράσεις της αγοράς, εστιάζοντας στα διαγράμματα και τους σχηματισμούς που δημιουργούνται σε αυτά.

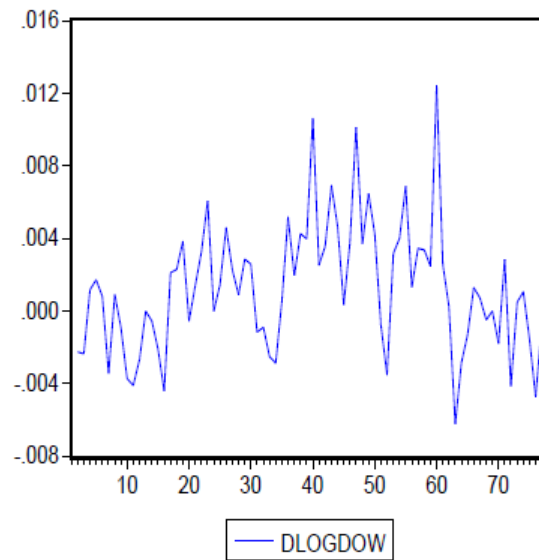
Η χρήση τόσο διαστρωματικών όσο και διαχρονικών δεδομένων για την αποτίμηση μετοχών είναι ευρέως διαδεδομένη καθώς επιτρέπει στους αναλυτές την εξαγωγή συμπερασμάτων με σχετικά ασφαλή τρόπο. Πριν περιγραφεί η διαδικασία εφαρμογής των panel δεδομένων στην περίπτωση του χρηματιστηρίου παρατίθενται ορισμένα διαγράμματα που αφορούν τη μεταβολή της τιμής των μετοχών:



Σχήμα 20: <sup>18</sup> Ημερήσιες τιμές κλεισίματος του δείκτη DowJones Utility index

<sup>18</sup> <http://www.learningmarkets.com/>

Σε συνέχεια του προηγούμενου διαγράμματος το επόμενο απεικονίζει το μετασχηματισμό των δεδομένων σε ημερήσιες αποδόσεις:

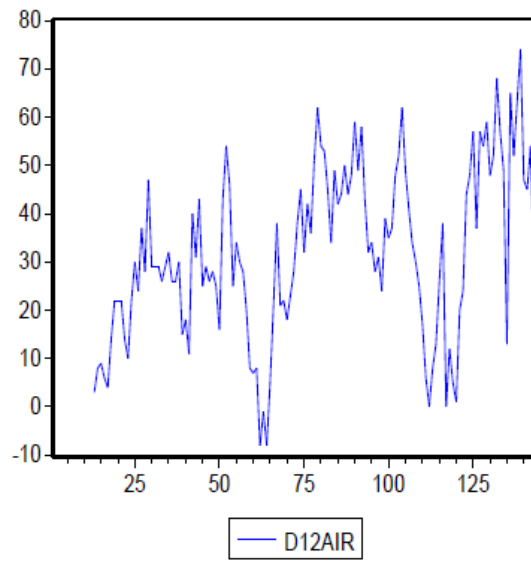


Σχήμα 21: Ημερήσιες τιμές κλεισίματος του δείκτη *DowJones Utility index (log)*<sup>19</sup>

Σε συνέχεια του προηγούμενου διαγράμματος το επόμενο απεικονίζει το μετασχηματισμό των δεδομένων σε ημερήσιες αποδόσεις:

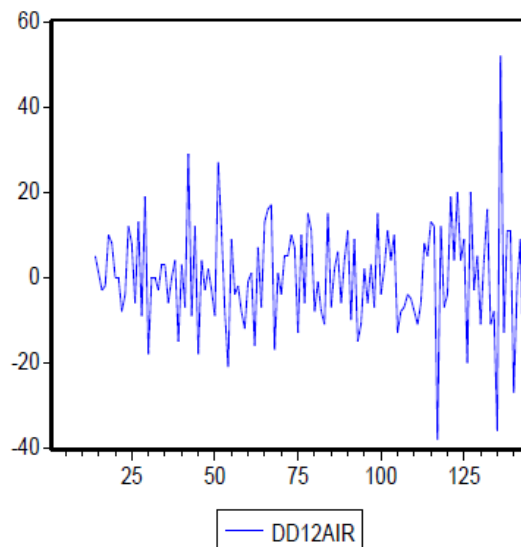
---

<sup>19</sup> <http://www.learningmarkets.com/>



Σχήμα 22: Μηνιαίες τιμές αεροπορικών επιβατών

Ένας ακόμη μετασχηματισμός μπορεί να συμβεί: Ο Μετασχηματισμός των δεδομένων στην πρώτη διάφορα των μηνιαίων διαφορών:



Σχήμα 23: Μηνιαίες τιμές αεροπορικών επιβατών - Μετασχηματισμός



### 4.3. ΠΑΡΑΔΕΙΓΜΑ - ΑΣΚΗΣΕΙΣ

Παρακάτω παρατίθενται ορισμένα παραδείγματα – ασκήσεις που αναδεικνύουν τη σπουδαιότητα των δεδομένων πάνελ. Τα παραδείγματα παραβάλλονται με τη μορφή ερωτοαπάντησης. Αναλυτικότερα:

#### ΠΑΡΑΔΕΙΓΜΑ 1

**Πώς μπορούσαν να βοηθήσουν τα στοιχεία «panel» στην περίπτωση που έχουμε:**

- Μεροληψία (εν δυνάμει) από παραλειπόμενες μεταβλητές που μεταβάλλονται διαστρωματικά, αλλά όχι διαχρονικά;
- Μεροληψία (εν δυνάμει) από παραλειπόμενες μεταβλητές που μεταβάλλονται διαχρονικά, αλλά όχι διαστρωματικά;

Να χρησιμοποιηθούν για την απάντηση τα αποτελέσματα από την εμπειρική ανάλυση στο σύνολο διαστρωματικών στοιχείων χρονολογικών σειρών («panel») από 48 πολιτείες των ΗΠΑ για 7 έτη (1982, ..., 1988) για την οδήγηση υπό την επήρεια αλκοόλ. Τα αποτελέσματα αυτά δίνονται στους δύο πιο κάτω πίνακες.

<b>Ανάλυση Παλινδρόμησης και το Αποτέλεσμα της Νομοθεσίας περί Οδήγησης υπό την Επήρεια Αλκοόλ στα Τροχαία Δυστυχήματα</b>						
<b>Εξαρτημένη μεταβλητή: Ποσοστό Τροχαίων Δυστυχημάτων (Θάνατοι ανά 10.000)</b>						
<b>Ερμηνευτική Μεταβλητή</b>	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>
<b>Φόρος Μπύρας</b>	0.36** (0.05)	-0.66** (0.20)	-0.64* (0.25)	-0.45* (0.22)	-0.70** (0.25)	-0.46** <b>(0.22)</b>
<b>Όριο ηλικίας νόμιμης κατανάλωσης αλκοόλ 18</b>				0.028 (0.066)	-0.011 (0.064)	
<b>Όριο ηλικίας νόμιμης κατανάλωσης αλκοόλ 19</b>				-0.019 (0.040)	-0.078 (0.049)	
<b>Όριο ηλικίας νόμιμης κατανάλωσης αλκοόλ 20</b>				0.031 (0.046)	-0.102* (0.046)	
<b>Ηλικία κατανάλωσης</b>						<b>-0.002</b>

αλκοόλ						<b>(0.017)</b>
Υποχρεωτική φυλάκιση				0.013 (0.032)	-0.026 (0.065)	
Υποχρεωτική θητεία σε δημόσιες υπηρεσίες				0.033 (0.115)	0.147 (0.137)	
Υποχρεωτική φυλάκιση ή υποχρεωτική θητεία σε δημόσιες υπηρεσίες						<b>0.031 (0.076)</b>
Μέσος όρος μιλίων οχημάτων ανά οδηγό				0.008 (0.008)	0.017 (0.010)	<b>0.009 (0.008)</b>
Ποσοστό ανεργίας				-0.063** (0.012)		<b>-0.063** (0.012)</b>
Πραγματικό εισόδημα κατά κεφαλήν (λογάριθμος)				1.81** (0.47)		<b>1.79** (0.45)</b>
Επιδράσεις στρώματος	Όχι	Ναι	Ναι	ναι	Ναι	<b>ναι</b>
Επιδράσεις χρόνου	Όχι	Όχι	Ναι	ναι	Ναι	<b>ναι</b>

Αυτές οι παλινδρομήσεις εκτιμήθηκαν με χρήση στοιχείων “panel” για 48 πολιτείες των ΗΠΑ από το 1982 έως το 1988 (συνολικά 336 παρατηρήσεις), που περιγράφονται στο παράρτημα 8.1. Τα τυπικά σφάλματα δίνονται στις παρενθέσεις κάτω από τους συντελεστές και οι πιθανότητες p δίνονται στις παρενθέσεις κάτω από τις τιμές της στατιστικής F. Οι ατομικοί συντελεστές είναι στατιστικά σημαντικοί σε επίπεδο \*5% και \*\*1%.

<b>Πίνακας Ανάλυση Παλινδρόμησης και το Αποτέλεσμα της Νομοθεσίας περί Οδήγησης υπό την 8.1 Επήρεια Αλκοόλ στα Τροχαία Δυστυχήματα</b>						
<b>Εξαρτημένη μεταβλητή: Ποσοστό Τροχαίων Δυστυχημάτων (Θάνατοι ανά 10.000)</b>						
	<b>(1)</b>	<b>(2)</b>	<b>(3)</b>	<b>(4)</b>	<b>(5)</b>	<b>(6)</b>
<b>Στατιστικές F και τιμές p για τον Έλεγχο Εξαίρεσης Ομάδων Μεταβλητών</b>						
<b>Επιδράσεις χρόνου = 0</b>			2.47	11.44	2.28	<b>11.59</b>

			(<0.001)	(<0.001)	(0.037)	(<0.001)
Συντελεστές νόμιμης ηλικίας κατανάλωσης αλκοόλ = 0				0.48 (0.696)	2.09 (0.102)	
Φυλάκιση, υποχρεωτική θητεία σε δημόσιες υπηρεσίες = 0				0.17 (0.845)	0.59 (0.557)	
Ποσοστό ανεργίας, εισόδημα κατά κεφαλήν = 0				38.29 (<0.001)		<b>40.12</b> (<0.001)
$\overline{R^2}$	0.090	0.889	0.891	0.926	0.893	<b>0.926</b>

Αυτές οι παλινδρομήσεις εκτιμήθηκαν με χρήση στοιχείων “panel” για 48 πολιτείες των ΗΠΑ από το 1982 έως το 1988 (συνολικά 336 παρατηρήσεις), που περιγράφονται στο παράρτημα 8.1. Τα τυπικά σφάλματα δίνονται στις παρενθέσεις κάτω από τους συντελεστές και οι πιθανότητες p δίνονται στις παρενθέσεις κάτω από τις τιμές της στατιστικής F. Οι ατομικοί συντελεστές είναι στατιστικά σημαντικοί σε επίπεδο \*5% και \*\*1%.

#### Απάντηση:

Στο υπόδειγμα 1 έχουμε μόνο τον φόρο μύρας ως ερμηνευτική μεταβλητή μεταβλητή. Παρατηρούμε ότι συντελεστής έχει θετικό πρόσημο(υποδηλώνοντας ότι υψηλότερος φόρος στην μύρα οδηγεί σε μεγαλύτερο αριθμό ατυχημάτων κάτι που δεν φαίνεται συμβατό με την οικονομική θεωρία) ενώ και η ερμηνευτική ικανότητα του υποδείγματος είναι χαμηλή.

Στο υπόδειγμα 2 έχουμε εισάγει και μία μεταβλητή  $\alpha_i$  η οποία μεταβάλλεται διαστρωματικά αλλά όχι διαχρονικά.

Δηλαδή το υπόδειγμα είναι:

$$\text{ποσοστό τροχαίων}_{it} = \beta_0 + \beta_1 \text{φόρος μύρας}_{it} + \alpha_i + \varepsilon_{it}$$

Η ερμηνευτική ικανότητα αυξάνεται σημαντικά και το πρόσημο του συντελεστή του φόρου είναι συμβατό με την οικονομικά θεωρία. Το  $\alpha_i$  περιλαμβάνει μεταβλητές που δεν περιλαμβάνονται στο υπόδειγμα ( ηλικία οχημάτων, ποιότητα δρόμων) οι οποίες θα μπορούσαν να προκαλούν μεροληψία από παραλειπόμενες μεταβλητές.

Στο υπόδειγμα 3 εισάγουμε και σταθερές επιδράσεις χρόνου,  $\mu_t$  για να ελέγξουμε για μεταβλητές που μεταβάλλονται χρονικά αλλά όχι διαστρωματικά (μεταβολές στην εθνική νομοθεσία αναφορικά με την οδήγηση και το ποτό) και δεν περιλαμβάνονται στο υπόδειγμα.

Το νέο υπόδειγμα είναι:

$$\text{ποσοστό τροχαίων}_{it} = \beta_0 + \beta_1 \text{φόρος μύρας}_{it} + \alpha_i + \mu_t + \varepsilon_{it}$$

Οι σταθερές επιδράσεις χρόνου είναι στατιστικά σημαντικές και η ερμηνευτική ικανότητα του υποδείγματος βελτιώνεται περισσότερο. Συνεπώς το υπόδειγμα 3 είναι καλύτερο.

## ΠΑΡΑΔΕΙΓΜΑ 2

Έστω το ακόλουθο υπόδειγμα:

$$W_{it} = \alpha_0 + \alpha_1 EDUC_i + \alpha_2 EXP_{it} + \mu_i + \varepsilon_{it}$$

Όπου:

- $W_{it}$ : είναι ο μισθός του ατόμου  $i$  το έτος  $t$
- $EDUC_i$ : είναι τα έτη σπουδών του ατόμου  $i$
- $EXP_{it}$ : είναι τα χρόνια προϋπηρεσίας του ατόμου  $i$
- $\mu_i$ : είναι ένας μη παρατηρήσιμος παράγοντας που μεταβάλλεται διαστρωματικά αλλά όχι διαχρονικά
- $\varepsilon_i$ : είναι ο διαταραχτικός όρος με μηδενικό μέσο και σταθερή διακύμανση

Θέλουμε να εκτιμήσουμε το υπόδειγμα σε ένα δείγμα 500 εργαζόμενων που απασχολούνται στον τομέα της σιδηρουργίας ηλικίας από 35 έως 45 χρόνων. Τα δεδομένα παρατηρούνται για μία περίοδο 5 χρόνων.

1. Ερμηνεύστε τους συντελεστές  $\alpha_1$ ,  $\alpha_2$ .
2. Δώστε μία ερμηνεία για την μη παρατηρήσιμη μεταβλητή  $\mu_i$
3. Ξαναγράψτε το υπόδειγμα σε πρώτες διαφορές
4. Αποφασίζουμε να εκτιμήσουμε το υπόδειγμα με την μέθοδο ελαχίστων τετραγώνων. Η μέθοδος αυτή θα οδηγούσε σε αμερόληπτους εκτιμητές?

### Απάντηση:

1.Ο συντελεστής  $\alpha_1$  δείχνει την επίδραση που έχουν τα έτη σπουδών στο μισθό εκκίνησης του ατόμου θεωρώντας όλους τους άλλους παράγοντες σταθερούς. Ο συντελεστής  $\alpha_2$  δείχνει την επίδραση στον μισθό από την αύξηση της προϋπηρεσίας κατά ένα χρόνο θεωρώντας όλους τους άλλους παράγοντες σταθερούς.

2.Η μη παρατηρήσιμη μεταβλητή  $\mu_i$  είναι η σταθερή επίδραση του στρώματος  $i$  και αναφέρεται σε παράγοντες όπως η προσπάθεια και η ικανότητα που είναι χαρακτηριστικά του ατόμου  $i$ .

3. Τα μοντέλο σε πρώτες διαφορές θα είναι:

$$W_{it} - W_{it-1} = \alpha_2(EXP_{it} - EXP_{it-1}) + \varepsilon_{it} - \varepsilon_{it-1}$$

Παρατηρούμε ότι γράφοντας το μοντέλο σε πρώτες διαφορές δεν μπορούμε να εκτιμήσουμε τον συντελεστή  $\alpha_1$  καθώς η μεταβλητή EDU δεν έχει χρονική διάσταση.

4-5: Η μέθοδος ελαχίστων τετραγώνων υπό την συνθήκη αυστηρής εξωγένειας:

$E(\varepsilon_{it} / X_{it}) = 0$  για κάθε  $t$  οδηγεί σε αμερόληπτους και συνεπείς εκτιμητές του υποδείγματος σε πρώτες διαφορές. Καθώς όμως τα νέα σφάλματα δεν είναι σφαιρικά ο εκτιμητής δεν είναι αποτελεσματικός. Αποδεικνύεται ότι αποτελεσματικός εκτιμητής είναι ο Generalized Least Squares estimator.

### ΠΑΡΑΔΕΙΓΜΑ 3

Έστω τα ακόλουθα δεδομένα:

Countryid	Year	Pdinc	Persexp	Hhldinc
Xylandia	1991	34000	25000	60000
Xylandia	1992	35000	26000	71000
Xylandia	1993	36050	26500	72000
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Xylandia	2001	45000	35000	92000
Bergunia	1991	23000	19000	55000
Bergunia	1992	24000	20100	57000
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Begunia	2001	40000	35000	88000
Taimat	1991	30999	20000	63000
Taimat	1992	31000	21000	62030
.	.	.	.	.
.	.	.	.	.
.	.	.	.	.
Taimat	2001	36000	28000	69040

Όπου

- pdinc: είναι το προσωπικό διαθέσιμο εισόδημα
- Persexp: είναι οι προσωπικές δαπάνες και
- Hhldinc: είναι το μέσο εισόδημα του νοικοκυριού.

1. Γράψτε το υπόδειγμα που συνδέει τις προσωπικές δαπάνες με το προσωπικό διαθέσιμο εισόδημα και το μέσο εισόδημα του νοικοκυριού.
2. Γράψτε το υπόδειγμα υποθέτοντας ότι υπάρχουν unit specific effects.
3. Γράψτε το υπόδειγμα υποθέτοντας ότι υπάρχουν temporal effects.
4. Πόσες παρατηρήσεις θα υπάρχουν για κάθε dummy variable? πόσες από αυτές θα είναι 0 και πόσες 1 στα ερωτήματα 2 και 3;

**Απάντηση:**

1. παρατηρούμε ότι έχουμε στοιχεία panel για τρεις χώρες σε 11 χρονικές περιόδους. Το υπόδειγμα που συνδέει τις τρεις μεταβλητές είναι:

$$Per\ exp_{it} = \beta_0 + \beta_1 pdinc_{it} + \beta_2 Hhldinc_{it} + u_{it}$$

2. Αν υποθέσουμε ότι υπάρχουν unit specific effects το υπόδειγμα θα περιλαμβάνει και έναν όρο  $\mu_i$ :

$$Per\ exp_{it} = \beta_0 + \mu_i + \beta_1 pdinc_{it} + \beta_2 Hhldinc_{it} + u_{it}$$

3. Αν υποθέσουμε ότι υπάρχουν temporal effects το υπόδειγμα θα περιλαμβάνει και έναν όρο  $\mu_t$ :

$$Per\ exp_{it} = \beta_0 + \mu_t + \beta_1 pdinc_{it} + \beta_2 Hhldinc_{it} + u_{it}$$

4. Αν θέλουμε να εκτιμήσουμε το υπόδειγμα 2 θα πρέπει να εισάγουμε σε αυτό 2 δυαδικές μεταβλητές ( γενικά  $n-1$  dummy variables):

$$Per\ exp_{it} = \beta_0 + \beta_1 pdinc_{it} + \beta_2 Hhldinc_{it} + \beta_3 country1 + \beta_4 country2 + u_{it}$$

Όπου

- η μεταβλητή country1 ορίζουμε να αναφέρεται στην χώρα xylandia ενώ
- η μεταβλητή country2 στη χώρα Bergunia.

Οι δύο δυαδικές μεταβλητές θα έχουν από 33 παρατηρήσεις όσες και κάθε άλλη μεταβλητή στο υπόδειγμα ( $3*11=33$ ).

Παρατήρηση	country1	country2
1	1	0
2	1	0
3	1	0
4	1	0
5	1	0
6	1	0
7	1	0
8	1	0
9	1	0
10	1	0
11	1	0
12	0	1
13	0	1
14	0	1
15	0	1
16	0	1
17	0	1
18	0	1
19	0	1
20	0	1
21	0	1
22	0	1
23	0	0
24	0	0
25	0	0
26	0	0
27	0	0
28	0	0
29	0	0
30	0	0
31	0	0
32	0	0
33	0	0

Οι πρώτες 11 παρατηρήσεις αφορούν την χώρα xylandia στα 11 χρόνια που την παρακολουθούμε, οι επόμενες 11 στην χώρα Bergunia ενώ οι τελευταίες 11 παρατηρήσεις των δύο μεταβλητών αφορούν την χώρα Taimat και είναι μηδενικές καθώς η χώρα αυτή δεν είναι ούτε η country1 ούτε η country2.

Αν τώρα θέλουμε να εκτιμήσουμε το υπόδειγμα του ερωτήματος 3 θα εισάγουμε 10 δυαδικές μεταβλητές ( $11-1=10$ ):



$$Per\ exp_{it} = \beta_0 + \beta_1 pdinc_{it} + \beta_2 Hhdinc_{it} + \beta_3 year\ 1 + \beta_4 year\ 2 + \beta_5 year\ 3 + \beta_6 year\ 4 + \beta_7 year\ 5 + \beta_8 year\ 6 + \beta_9 year\ 7 + \beta_{10} year\ 8 + \beta_{11} year\ 9 + \beta_{12} year\ 10 + u_{it}$$

Όπου *year1* έως *year10* οι χρονιές από το 1991 έως το 2000 αντίστοιχα. Δηλαδή θεωρούμε ότι μόνο η μεταβλητή 2001 δεν αντιπροσωπεύεται από κάποια δυαδική μεταβλητή. Κάθε μεταβλητή θα έχει 33 παρατηρήσεις που θα παίρνουν τιμές 0 και 1:

	Year1	Year2	Year3	Year4	Year5	Year6	Year7	Year8	Year9	Year10
1	1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	0	0
3	0	0	1	0	0	0	0	0	0	0
4	0	0	0	1	0	0	0	0	0	0
5	0	0	0	0	1	0	0	0	0	0
6	0	0	0	0	0	1	0	0	0	0
7	0	0	0	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	0	0	1	0
10	0	0	0	0	0	0	0	0	0	1
11	0	0	0	0	0	0	0	0	0	0
12	1	0	0	0	0	0	0	0	0	0
13	0	1	0	0	0	0	0	0	0	0
14	0	0	1	0	0	0	0	0	0	0
15	0	0	0	1	0	0	0	0	0	0
16	0	0	0	0	1	0	0	0	0	0
17	0	0	0	0	0	1	0	0	0	0
18	0	0	0	0	0	0	1	0	0	0
19	0	0	0	0	0	0	0	1	0	0
20	0	0	0	0	0	0	0	0	1	0
21	0	0	0	0	0	0	0	0	0	1
22	0	0	0	0	0	0	0	0	0	0
23	1	0	0	0	0	0	0	0	0	0
24	0	1	0	0	0	0	0	0	0	0
25	0	0	1	0	0	0	0	0	0	0
26	0	0	0	1	0	0	0	0	0	0
27	0	0	0	0	1	0	0	0	0	0
28	0	0	0	0	0	1	0	0	0	0
29	0	0	0	0	0	0	1	0	0	0
30	0	0	0	0	0	0	0	1	0	0
31	0	0	0	0	0	0	0	0	1	0

32	0	0	0	0	0	0	0	0	0	1
33	0	0	0	0	0	0	0	0	0	0

Παρατηρούμε ότι η σειράς 11 (για την χώρα xylandia ), 22 (για την χώρα Bergunia) και 33 (για την χώρα Taimat) αναφέρονται στο 2001 και συνεπώς έχουν τιμή 0 για όλες τις δυαδικές μεταβλητές.

#### ΠΑΡΑΔΕΙΓΜΑ 4

Θεωρούμε το ακόλουθο υπόδειγμα κατανάλωσης αλκοολούχων ποτών για 10 άτομα σε 10 χρονικές περιόδους:

$$\log(C_{it}) = \beta_0 + \beta_1 \log(R_{it}) + \beta_2 \log(P_{it}) + \lambda_i + u_{it}$$

Όπου:

- $C_{it}$  είναι το επίπεδο κατανάλωσης αλκοολούχων ποτών
- $R_{it}$  είναι το επίπεδο εισοδήματος του ατόμου
- $P_{it}$  είναι η τιμή των ποτών και
- $\lambda_i$  είναι ένας παράγοντας μη παρατηρήσιμος.

1. Δώστε την οικονομική ερμηνεία των συντελεστών  $\beta_1$  και  $\beta_2$
2. Δώστε μία ερμηνεία για την μεταβλητή  $\lambda_i$ . Τι πρόβλημα δημιουργεί η ύπαρξη της στο υπόδειγμα?
3. Γράψτε το υπόδειγμα σε πρώτες διαφορές? Σχολιάστε την συνέπεια και αποτελεσματικότητα των εκτιμητών. Ποια άλλη μέθοδος θα μας επέτρεπε την εκτίμηση του υποδείγματος;

**Απάντηση:**

1. οι συντελεστές  $\beta_1$  και  $\beta_2$  εκφράζουν την εισοδηματική ελαστικότητα και ελαστικότητα τιμής αντίστοιχα καθώς το υπόδειγμα είναι εκφρασμένο σε λογαρίθμους. Από την μικροοικονομική γνωρίζουμε ότι η ελαστικότητα τιμής δίνεται από τον ακόλουθο τύπο:

$$\varepsilon_{PC} = \frac{dC}{dP} \times \frac{P}{C}$$

Ο συντελεστής  $\beta_2$  στο υπόδειγμα είναι η παράγωγος του  $\log C$  ως προς  $\log P$ .

Αρκεί άρα να δείξουμε ότι η παράγωγος αυτή ισούται με την ελαστικότητα.

$$\beta_2 = \frac{d \log C(P)}{d \log P}$$

$$\frac{d \log C(P)}{dP} = \frac{1}{C} \times \frac{dC}{dP}$$

$$\frac{d \log P}{dP} = \frac{1}{P}$$

Και αντικαθιστώντας στον τύπο του  $\beta_2$ :

$$\beta_2 = \frac{\frac{1}{C} \times \frac{dC}{dP} \times dP}{\frac{1}{P} \times dP} = \frac{P}{C} \times \frac{dC}{dP} = \varepsilon_{PC}$$

Ομοίως αποδεικνύεται ότι το  $\beta_1$  είναι η εισοδηματική ελαστικότητα.

2. Η μεταβλητή  $\lambda$  είναι ένας παράγοντας που επιδρά στην ποσότητα κατανάλωσης αλκοόλ, δεν είναι παρατηρήσιμος και μεταβάλλεται διαστρωματικά αλλά όχι διαχρονικά. Θα μπορούσε να συνδέεται με τις προτιμήσεις του ατόμου. Η μεταβλητή αυτή ωστόσο δεν παρατηρείται και η ύπαρξή της στο υπόδειγμα θα μπορούσε να δημιουργεί πρόβλημα ενδογένειας αν αυτή συνδεόταν με το διαταραχτικό όρο του υποδείγματος.

3. Το υπόδειγμα μπορεί να εκτιμηθεί γράφοντας το σε πρώτες διαφορές είτε χρησιμοποιώντας δυαδικές μεταβλητές στρώματος (στην συγκεκριμένη άσκηση 9). Με

τις δύο αυτές μεθόδους εξαλείφουμε τον παράγοντα  $\lambda$  που δεν είναι παρατηρήσιμος και οι εκτιμητές που παίρνουμε είναι συνεπείς και αμερόληπτοι. Το πρόβλημα είναι ότι οι εκτιμητές δεν είναι αποτελεσματικοί (δεν έχουν δηλαδή την μικρότερη δυνατή διακύμανση) κάτι που θα μας εξασφάλιζε υπό ορισμένες προϋποθέσεις η μέθοδος των γενικευμένων ελαχίστων τετραγώνων.

Το υπόδειγμα σε πρώτες διαφορές είναι το ακόλουθο:

$$\log(C_{it}) - \log(C_{it-1}) = \beta_1(\log(R_{it}) - \log(R_{it-1})) + \beta_2(\log(P_{it}) - \log(P_{it-1})) + u_{it} - u_{it-1}$$

$$\Leftrightarrow \log\left(\frac{C_{it}}{C_{it-1}}\right) = \beta_1 \log\left(\frac{R_{it}}{R_{it-1}}\right) + \beta_2 \log\left(\frac{P_{it}}{P_{it-1}}\right) + u_{it} - u_{it-1}$$

### ΠΑΡΑΔΕΙΓΜΑ 5

Έστω το εξής εκτιμημένο υπόδειγμα με στοιχεία panel για 10 στρώματα (έστω χώρες)  $i = 1, 2, 3, \dots, 10$  παρατηρούμενα για 100 χρόνια ( $t = 0, \dots, 99$ )

$$y_{it} = 10.23 + 0.5D_{1t} + 1.2D_{2t} + \dots + 7D_{9t} - 5.3x_{it} + \hat{e}_{it}$$

όπου  $D$  οι δυαδικές στρώματος (χώρας) και  $e$  ο διαταρακτικός όρος.

(i) Πόσες παρατηρήσεις έχουμε για τη  $D_{2t}$ ; Ποια είναι η τιμή της  $D_{1t}$  και  $D_{3t}$

-για το στρώμα (χώρα) 1 το χρόνο 3

-για το στρώμα (χώρα) 1 το χρόνο 15

-για το στρώμα (χώρα) 3 το χρόνο 0

-για το στρώμα (χώρα) 9 το χρόνο  $t$

(ii) Ερμηνεύστε τους συντελεστές των ψευδομεταβλητών  $D_{1t}$  και  $D_{9t}$

(iii) Ποιου είδους μεροληψία εξαλείφεται με τη χρήση του συγκεκριμένου υποδείγματος και στοιχείων; Ποια ενδέχεται να παραμένει;

(iv) Εκφράστε το εκτιμημένο υπόδειγμα σε όρους αποκλίσεων από τους διαχρονικούς και διαστρωματικούς μέσους.

### Απάντηση

(i) Όπως κάθε μεταβλητή στο υπόδειγμα έτσι και η  $D_{2t}$  θα έχει παρατηρήσεις για 100 χρόνια 10 χωρών, συνολικά 1000.

Αντίστοιχα για  $D_{1t}$  και  $D_{3t}$  οι τιμές θα είναι:

- 1 και 0

- 1 και 0

- 0 και 1

- 0 και 0

Δηλαδή οι 1000 παρατηρήσεις κάθε μιας δυαδικής χώρας θα είναι μηδενικά εκτός από τις παρατηρήσεις που αναλογούν στην αντίστοιχη χώρα, όπου θα παίρνουν την τιμή 1.

(ii) Η προβλεπόμενη τιμή της  $y$  από το υπόδειγμα (τη συμβολίζουμε με καπελάκι) είναι:

$$\hat{y}_{it} = 10.23 + 0.5D_{1t} + 1.2D_{2t} + \dots + 7D_{9t} - 5.3x_{it}$$

Για  $x_{it} = 0$ :

$$\hat{y}_{it} |_{x_{it}=0} = 10.23 + 0.5D_{1t} + 1.2D_{2t} + \dots + 7D_{9t} \quad (*)$$

Για το στρώμα 10 η (\*) γίνεται:

$$\hat{y}_{10t} |_{x_{it}=0} = 10.23 \quad (1)$$

Για το στρώμα 1 η (\*) γίνεται:

$$\hat{y}_{1t} |_{x_{it}=0} = 10.23 + 0.5 \quad (2)$$

Αφαιρώντας την (2) από την (1) έχουμε:

$$\hat{y}_{1t} |_{x_{it}=0} - \hat{y}_{10t} |_{x_{it}=0} = 0.5$$

και επομένως ο συντελεστής της  $D_{1t}$  ερμηνεύεται ως εξής: Η προβλεπόμενη τιμή της  $y$  του στρώματος 1 για  $x_{it} = 0$  είναι μεγαλύτερη κατά 0.5 μονάδες μέτρησης της  $y$  από αυτή του στρώματος 10.

Όμοια για τον συντελεστή του  $D_{9t}$ , για το στρώμα 9 η (\*) γίνεται:

$$\hat{y}_{9t} |_{x_{it}=0} = 10.23 + 7 \quad (3)$$

Αφαιρώντας την (3) από την (1) έχουμε:

$$\hat{y}_{9t} |_{x_{it}=0} - \hat{y}_{10t} |_{x_{it}=0} = 7$$

και επομένως ο συντελεστής της  $D_{9t}$  ερμηνεύεται ως εξής: Η προβλεπόμενη τιμή της  $y$  του στρώματος 9 για  $x_{it} = 0$  είναι μεγαλύτερη κατά 7 μονάδες μέτρησης της  $y$  από αυτή του στρώματος 10.

iii) Το υπόδειγμα χρησιμοποιεί panel data (χρονολογικής και διαστρωματικής διάστασης). Γίνεται εκτίμηση με τη μέθοδο των n-1 δυαδικών μεταβλητών που απομονώνουν την επίδραση του στρώματος. Κατ'αυτόν τον τρόπο μπορούμε να συμπεριλάβουμε όλους τους μη μετρήσιμους παράγοντες που επηρεάζουν την  $y$  πέραν της  $x$  και οι οποίοι δεν μεταβάλλονται διαχρονικά. Οι παράγοντες αυτοί ενδέχεται να συσχετίζονται με το  $x$  διαστρωματικά και η παράλειψή τους θα καθιστούσε το  $x$  ενδογενές και τους OLS εκτιμητές μεροληπτικούς και ασυνεπείς.

Το πρόβλημα παραμένει αν υπάρχουν παράγοντες που συσχετίζονται με το  $x$  διαστρωματικά και μεταβάλλονται και διαχρονικά, οπότε η μέθοδος αυτή δεν μπορεί να τους ενσωματώσει (χρειάζονται τουλάχιστον και δυαδικές μεταβλητές χρόνου).

iv) Ο διαχρονικός μέσος είναι για κάθε στρώμα ο μέσος όρος των παρατηρήσεων όλων των χρόνων. Θα έχουμε επομένως 10 διαχρονικούς μέσους, όσα και τα στρώματα. Ο διαστρωματικός μέσος είναι για κάθε χρόνο ο μέσος όρος των παρατηρήσεων όλων των στρωμάτων. Θα έχουμε επομένως 100 διαστρωματικούς μέσους.

Για το στρώμα 1 η προβλεπόμενη τιμή της  $y$  από την εξειδίκευση της εκφώνησης:

$$\hat{y}_{1t} = 10.23 + 0.5 - 5.3x_{1t}$$

Επομένως ο διαχρονικός της μέσος είναι:

$$\bar{\hat{y}}_1 = \frac{1}{100} \sum_{t=0}^{T=99} \hat{y}_{1t} = 10.23 + 0.5 - 5.3 \frac{1}{100} \sum_{t=0}^{T=99} x_{1t}$$

Αρα η προβλεπόμενη αποκλιση του στρώματος 1 από το διαχρονικό του μέσο είναι:

$$y_{1t} - \bar{\hat{y}}_1 = -5.3(x_{1t} - \bar{x}_1)$$

Με παρόμοιο τρόπο μπορούμε να πούμε γενικά για το στρώμα  $i$  :

$$y_{it} - \bar{\hat{y}}_i = -5.3(x_{it} - \bar{x}_i)$$

Για το χρόνο 1 η προβλεπόμενη τιμή της  $y$  από την εξειδίκευση της εκφώνησης:

$$\hat{y}_{i1} = 10.23 + 0.5D_{11} + 1.2D_{21} + \dots + 7D_{91} - 5.3x_{i1}$$

Επομένως ο διαστρωματικός της μέσος είναι:

$$\bar{\hat{y}}_1 = \frac{1}{10} \sum_{i=1}^{N=10} \hat{y}_{i1} = 10.23 + 0.5 \frac{1}{10} \sum_{i=1}^{N=10} D_{11} + 1.2 \frac{1}{10} \sum_{i=1}^{N=10} D_{21} + \dots + 7 \frac{1}{10} \sum_{i=1}^{N=10} D_{91} - 5.3 \frac{1}{10} \sum_{i=1}^{N=10} x_{i1}$$

Αλλα τα αθροίσματα των δυαδικών μεταβλητών ισούνται με τη μονάδα εφόσον οι δυαδικές μεταβλητές για τον χρόνο 1 παίρνουν παντού τιμές ίσες με το μηδέν εκτός από το στρώμα στο οποίο αντιστοιχούν. Επομένως:

$$\bar{\hat{y}}_1 = 10.23 + 0.5 \frac{1}{10} + 1.2 \frac{1}{10} + \dots + 7 \frac{1}{10} - 5.3 \frac{1}{10} \sum_{i=1}^{N=10} x_{i1}$$

Άρα η προβλεπόμενη απόκλιση του χρόνου 1 από το διαστρωματικό του μέσο είναι

$$\hat{y}_{i1} - \bar{\hat{y}}_1 = 0.5 \left( D_{11} - \frac{1}{10} \right) + 1.2 \left( D_{21} - \frac{1}{10} \right) + \dots + 7 \left( D_{91} - \frac{1}{10} \right) - 5.3 (x_{i1} - \bar{x}_1)$$

Με παρόμοιο τρόπο μπορούμε να πούμε γενικά ότι για τον χρόνο  $t$ :

$$\hat{y}_{it} - \bar{\hat{y}}_t = 0.5 \left( D_{1t} - \frac{1}{10} \right) + 1.2 \left( D_{2t} - \frac{1}{10} \right) + \dots + 7 \left( D_{9t} - \frac{1}{10} \right) - 5.3 (x_{it} - \bar{x}_t)$$

#### 4.4. ΤΟ ΟΙΚΟΝΟΜΕΤΡΙΚΟ ΠΑΚΕΤΟ EVIEWS

Πολλά λογισμικά πακέτα έχουν δημιουργηθεί για την εμπειρική ανάλυση και την ερμηνεία των οικονομετρικών υποδειγμάτων. Η εξέλιξη της τεχνολογίας εξάλλου είναι ραγδαία, με αποτέλεσμα το κομμάτι της οικονομικής επιστήμης να μην μείνει ανεπηρέαστο. Ένα πολύ χρήσιμο και εύχρηστο λογισμικό πακέτο ανάλυσης, βάση του οποίου θα παρουσιάζεται η εφαρμογή της θεωρητικής βάσης της Οικονομετρίας είναι το EVIEWS.

Το EVIEWS είναι ένα από τα πιο γνωστά προγράμματα υπολογιστών που χρησιμοποιούνται για την επίλυση οικονομετρικών προβλημάτων. Ένα από τα κυριότερα χαρακτηριστικά αυτού είναι η ευκολία στη χρήση και το φιλικό περιβάλλον διαχείρισης.

Το EVIEWS είναι ένα πρόγραμμα που παρέχει τη δυνατότητα στατιστικής και οικονομετρικής ανάλυσης στοιχείων βασισμένο στο παραθυρικό περιβάλλον εργασίας των WINDOWS.

Με το EVIEWS μπορεί κάποιος πολύ γρήγορα και απλά:

- Να καταχωρήσει τις μεταβλητές που πρόκειται να χρησιμοποιήσει και με βάση αυτές να δημιουργήσει νέες μεταβλητές

- Να αναλύσει στατιστικά τις μεταβλητές αυτές και
- Να εκτιμήσει οικονομετρικά σχέσεις μεταξύ των μεταβλητών χρησιμοποιώντας διάφορες παραδοσιακές αλλά και σύγχρονες οικονομετρικές μεθόδους.

Το πακέτο EViews περιλαμβάνει μια μεγάλη σειρά τεχνικών εκτίμησης οικονομικών υποδειγμάτων (μέθοδος ελαχίστων τετραγώνων, μέθοδοι για συστήματα, μη γραμμικά και δυναμικά υποδείγματα κοκ), μεθόδους για προβλέψεις και γραφικά, μη παραμετρικές μεθόδους για την εκτίμηση συναρτήσεων πυκνότητας πιθανότητας κλπ. Χρησιμοποιείται κυρίως όταν θέλουμε να αναλύσουμε οικονομικές χρονολογικές σειρές σε ένα επίπεδο ανάλυσης που υπερβαίνει την απλή στατιστική ανάλυση και προσεγγίζει περισσότερο την σύγκριση και εκτίμηση εναλλακτικών υποδειγμάτων, την διεξαγωγή προβλέψεων με εναλλακτικά υποδείγματα, την διεξαγωγή πολύπλοκων οικονομετρικών ελέγχων (πχ σταθερότητα του υποδείγματος, έλεγχοι για αυτοσυσχέτιση και λανθασμένη εξειδίκευση, έλεγχοι περιορισμών που θέτει η οικονομική θεωρία στις παραμέτρους κλπ).



## ΣΥΜΠΕΡΑΣΜΑΤΑ

Τα οικονομικά φαινόμενα είναι κατά βάση στοχαστικά, δηλαδή περιλαμβάνουν τυχαίους παράγοντες που αγνοούνται από τη μαθηματική λογική. Η οικονομετρία έρχεται να καλύψει το κενό ανάμεσα στις προσδιοριστικές σχέσεις και την οικονομική πραγματικότητα βάση της θεωρητικής στοχαστικής θεώρησης των οικονομικών σχέσεων.

Η έννοια της οικονομετρίας και η χρησιμότητα αυτής είναι εμφανής σε όλη την έκταση της παρούσας εργασίας. Είτε χρονολογική είτε διαστρωματική είναι η ανάλυση που πραγματοποιείται τα οφέλη είναι σημαντικά χωρίς αυτό να σημαίνει ότι σε κάθε μία από αυτές τις αναλύσεις δεν εμπεριέχονται και μειονεκτήματα.

Ο συνδυασμός των δυο αυτών τεχνικών, δηλαδή η διαστρωματική ανάλυση χρονολογικών σειρών ή η χρονολογική ανάλυση διαστρωματικών δεδομένων, ή με άλλα λόγια η ανάλυση δεδομένων πάνελ είναι ιδιαίτερα σημαντική και τα αποτελέσματα αυτής ιδιαίτερα ωφέλιμα για τους αναλυτές πολλών κλάδων.

Σε κάθε περίπτωση, προκειμένου μια ανάλυση να πραγματοποιηθεί και να οδηγήσει σε αναγνώσιμα και ερμηνεύσιμα αποτελέσματα χρησιμοποιείται η έννοια της στατιστικής και η εφαρμογή αυτής με τη βοήθεια του οικονομετρικού/ στατιστικού πακέτου EVIEWS. Η ναυτιλία, το χρηματιστήριο, η ζήτηση, η προσφορά, η μακροοικονομική σύγκλιση οικονομιών και άλλοι τομείς μπορούν να προσεγγιστούν και να ερμηνευθούν αξιοποιώντας κατάλληλα τη χρήση των δεδομένων πάνελ.

## **ΒΙΒΛΙΟΓΡΑΦΙΑ**

Ανδρικόπουλος, Α. (2000). Οικονομετρία. Θεωρία και Εμπειρικές Εφαρμογές. Εκδόσεις Μπένου, Αθήνα.

Βάμβουκας, Γ. (2007). Σύγχρονη Οικονομετρία, Ανάλυση και Εφαρμογές. Εκδόσεις Οικονομικού Πανεπιστημίου Αθηνών.

Δημέλη. Σ. (2013). Σύγχρονες Μέθοδοι Ανάλυσης Χρονολογικών Σειρών, Εκδόσεις Οικονομικού Πανεπιστημίου Αθηνών, Αθήνα.

Δριτσάκη, Χ. και Μ. Δριτσάκη (2013.) Εισαγωγή στην Οικονομετρία με τη χρήση του λογισμικού EViews, Εκδόσεις Κλειδάριθμος, Αθήνα.

Κάτος, Α. (2004). Οικονομετρία, Θεωρία και Εφαρμογές. Εκδόσεις Ζυγός.

Κιντής, Α. (2010). Σύγχρονη Οικονομετρική Ανάλυση. Εκδόσεις Gutenberg.

Συριόπουλος, Κ, Φίλιππας, Δ. (2010). Οικονομετρικά Υποδείγματα, Εκδόσεις Ανικούλα, Θεσσαλονίκη.

Τζαβαλής. Η. (2008). Οικονομετρία, Εκδόσεις Οικονομικού Πανεπιστημίου Αθηνών, Αθήνα

Χάλκος Γ., (2011) «Οικονομετρία», Εκδόσεις Gutenberg, Αθήνα

Χρήστου, Κ.Γ. (2002). Εισαγωγή στην Οικονομετρία. Εκδόσεις Gutenberg, Αθήνα.

Baltagi, B.H. (2008). Econometric Analysis of Panel Data, John Wiley & Sons, Ltd. U.K.

Dougherty C. (2006). Introduction to Econometrics (3rd Ed.), Oxford University Press. ISBN: 9780199280964.

Dritsakis, N. (2011). “Demand for money in Hungary: An ARDL Approach”, Review of Economics and Finance, Vol. 1, No. 5, pp 1-16.

Fisher, I (1937). Note on a short-cut method for calculating distributed lags, International Statistical Bulletin, Vol. 29 pp. 323 - 328.

Goldberger, A. S. (1991). A Course in Econometrics. Cambridge, MA: Harvard University Press. ISBN: 9780674175440.

Hamilton, J. D. (1994). Time Series Analysis, New Jersey, Princeton University Press.

Hondroyannis, G. (2000). The wage growth and inflation nexus in a dynamic multivariate context: New evidence from Greece. International Economic Journal, Vol.17, No.1, pp. 121 – 138.

Johnston J. and Dinardo J. (2005). Οικονομετρικές Μέθοδοι, 4η Αμερικάνικη Έκδοση, Εκδόσεις Κλειδάριθμος.

Maddala G.S. (1992). Introductory Econometrics, New Jersey, Prentice-Hall.

Myers R. (1990). Classical and Modern Regression with Applications, Belmont California, Duxbury Press.

Thomas R.L. (1997). Modern Econometrics: An Introduction. Harlow, Addison-Wesley

Wooldridge, J., (2006). Εισαγωγή στην Οικονομετρία, Μια Νέα Προσέγγιση. Τόμος Α, Εκδόσεις Παπαζήση.

## **ΗΛΕΚΤΡΟΝΙΚΕΣ ΠΗΓΕΣ**

<http://www.ssc.wisc.edu/~bhansen/>

<http://www.economicsnetwork.ac.uk/>

<http://slideplayer.gr/slide/1900726/>

[http://androulakis.bma.upatras.gr/mediawiki/index.php?title=%CE%A3%CE%91%CE%A3%CE%A3%CE%91%CE%9B%CE%9F%CE%A5\\_%CE%95%CE%A5%CE%A6%CE%A1%CE%9F%CE%A3%CE%A5%CE%9D%CE%97&printable=yes](http://androulakis.bma.upatras.gr/mediawiki/index.php?title=%CE%A3%CE%91%CE%A3%CE%A3%CE%91%CE%9B%CE%9F%CE%A5_%CE%95%CE%A5%CE%A6%CE%A1%CE%9F%CE%A3%CE%A5%CE%9D%CE%97&printable=yes)

<http://mathimataoikonometrias.blogspot.gr/2012/02/blog-post.html>

<https://www.google.gr/url?sa=i&rct=j&q=&esrc=s&source=images&cd=&cad=rja&uact=8&ved=0CAYQjB0&url=http%3A%2F%2Fstavrakoudis.econ.uoi.gr%2Fstavrakoudis%2F%3Fiid%3D204&ei=bueEVPiABMn8UtzlgNgP&bvm=bv.80642063,d.d24&psig=AFQjCNFJwI3sj-5kkBOrjpK45heQu6SBRQ&ust=1418082534910173>

<http://slideplayer.gr/slide/1987676/>

<http://pubs.sciepub.com/ajams/2/1/4/>

<https://onlinecourses.science.psu.edu/stat510/node/60>

<ftp://filer.soc.uoc.gr/>

[www.clarkson.com](http://www.clarkson.com)

[www.euro2day.gr](http://www.euro2day.gr)

[digilib.lib.unipi.gr/dspace/bitstream/unipi/5188/1/](http://digilib.lib.unipi.gr/dspace/bitstream/unipi/5188/1/)

<http://www.learningmarkets.com/>