

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ
ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ
ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

ΜΕΛΕΤΗ ΚΑΙ ΕΚΤΙΜΗΣΗ ΥΠΟΔΕΙΓΜΑΤΩΝ ΜΕ ΤΗ ΜΕΘΟΔΟ ΔΙΑΦΟΡΩΝ ΣΕ ΔΙΑΦΟΡΕΣ

Παναγιώτης Π. Ξανθόπουλος

Διπλωματική Εργασία

που υποβλήθηκε στο Τμήμα Στατιστικής και Ασφαλιστικής
Επιστήμης του Πανεπιστημίου Πειραιώς ως μέρος των
απαιτήσεων για την απόκτηση του Μεταπτυχιακού
Διπλώματος Ειδίκευσης στην Εφαρμοσμένη Στατιστική

Πειραιάς
Νοέμβριος 2014

Η παρούσα Διπλωματική Εργασία εγκρίθηκε ομόφωνα από την Τριμελή Εξεταστική Επιτροπή που ορίστηκε από τη ΓΣΕΣ του Τμήματος Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς στην υπ' αριθμ. συνεδρίασή του σύμφωνα με τον Εσωτερικό Κανονισμό Λειτουργίας του Προγράμματος Μεταπτυχιακών Σπουδών στην Εφαρμοσμένη Στατιστική

Τα μέλη της Επιτροπής ήταν:

- Καθ. Αγιακλόγλου Χρήστος (Επιβλέπων)
- Καθ. Γκλεζάκος Μιχαήλ
- Καθ. Τσίμπος Κλέων

Η έγκριση της Διπλωματική Εργασίας από το Τμήμα Στατιστικής και Ασφαλιστικής Επιστήμης του Πανεπιστημίου Πειραιώς δεν υποδηλώνει αποδοχή των γνώμων του συγγραφέα.

UNIVERSITY OF PIRAEUS



**DEPARTMENT OF STATISTICS
AND INSURANCE SCIENCE**

**POSTGRADUATE PROGRAM IN
APPLIED STATISTICS**

**STUDY AND ESTIMATION MODELS
USING THE DIFFERENCE IN
DIFFERENCES METHOD**

By

Panagiotis P. Ksanthopoulos

MSc Dissertation

submitted to the Department of Statistics and Insurance
Science of the University of Piraeus in partial fulfilment of the
requirements for the degree of Master of Science in Applied
Statistics

Piraeus, Greece
November 2014

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς

*Στους γονείς μου
Πάλο και Ξένη,
στον αδερφό μου Πέτρο
και στην Αναστασία*

Πανεπιστήμιο Πειραιώς

Ευχαριστίες

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά τον κ. Αγιακλόγλου Χρήστο, Πρόεδρο και Καθηγητή του Τμήματος Οικονομικής Επιστήμης του Πανεπιστημίου Πειραιώς για την αμέριστη βοήθεια και καθοδήγησή του, σε όλα τα στάδια συγγραφής της Διπλωματικής Εργασίας. Επίσης, θα ήθελα να ευχαριστήσω όλους τους καθηγητές του Μεταπτυχιακού Προγράμματος Εφαρμοσμένης Στατιστικής για τις πολύτιμες γνώσεις και τα εφόδια που μου προσέφεραν. Τέλος, θα ήθελα να ευχαριστήσω την οικογένεια μου και την Αναστασία, που βρίσκονταν καθ' όλη τη διάρκεια αυτής της προσπάθειας δίπλα μου, προσφέροντας μου τα απαραίτητα ψυχικά εφόδια.

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς

Περίληψη

Η εργασία αυτή έχει ως αντικείμενο μελέτης την επίδραση της Ευρωπαϊκής χρηματοπιστωτικής κρίσης στις πωλήσεις των πετρελαιοκίνητων οχημάτων, για τις χώρες που ανήκουν στην Ευρωπαϊκή Ένωση των 15 κρατών μελών και στην Ευρωπαϊκή Ζώνη Ελεύθερων Συναλλαγών ΕFTA 3. Η μέθοδος που χρησιμοποιείται είναι αυτή των διαφορών σε διαφορές για δεδομένα τύπου πάνελ, θεωρώντας ως ομάδα ελέγχου το πλήθος των συνολικών πωλήσεων των βενζινοκίνητων οχημάτων και ως ομάδα αντιμετώπισης το πλήθος των συνολικών πωλήσεων των πετρελαιοκίνητων οχημάτων στις παραπάνω χώρες, για τα έτη 2010 και 2011.

Η απαλοιφή των παρατήρητων επιδράσεων για κάθε χώρα πραγματοποιείται με τη λήψη των διαφορών μεταξύ των δύο περιόδων. Μέσω του κατάλληλου υποδείγματος παλινδρόμησης προκύπτει ότι η χρηματοπιστωτική κρίση επέφερε στατιστικά σημαντική μεταβολή στις πωλήσεις των πετρελαιοκίνητων οχημάτων για το έτος 2011, ίση με 55.621 οχήματα.

Πανεπιστήμιο Πειραιώς

Abstract

The study examines the impact of the European financial crisis on sales of diesel vehicles for countries belonging to the European Union of 15 and the European Free Trade Association, EFTA 3. The method used is the Difference in Differences for panel data, considering as control group the number of total sales of petrol cars and as treatment group the number of total sales of diesel vehicles in the above countries, for the years 2010 and 2011.

The deletion of the unobserved effects for each country is carried out by taking the differences between the two periods. Through the appropriate regression model shows that the financial crisis has led to significant changes in the sales of diesel vehicles in the year 2011, equal to 55.621 vehicles.

Πανεπιστήμιο Πειραιώς

Περιεχόμενα

Κατάλογος Πινάκων	xv
Κατάλογος Διαγραμμάτων	xvii
1. Βασικές Οικονομετρικές Έννοιες και Μέθοδοι	1
1.1 Εισαγωγή	1
1.2 Κατηγορίες Δεδομένων	2
1.3 Ανάλυση Παλινδρόμησης	6
1.4 Χρονοσειρές	12
1.5 Δεδομένα Πάνελ	20
1.6 Ανακεφαλαίωση	29
2. Μέθοδος Διαφορών σε Διαφορές	30
2.1 Εισαγωγή	30
2.2 Προβλήματα στην Εγκυρότητα	31
2.3 Εκτιμητές Απλών Διαφορών	32
2.4 Εκτιμητής Διαφορών σε Διαφορές	37
2.5 Αμεροληψία Εκτιμητών Απλών Διαφορών	40
2.6 Υποθέσεις για το Υπόδειγμα Διαφορών σε Διαφορές	41
2.7 Προεκτάσεις του Υποδείγματος Διαφορές σε Διαφορές	44
2.8 Ελέγχοντας την Εσωτερική Εγκυρότητα	46
2.9 Ανακεφαλαίωση	48
3. Αναφορά σε Εμπειρικές Μελέτες	49
3.1 Εισαγωγή	49
3.2 Κατώτατος Μισθός και Απασχόληση στις Γυναίκες του Oregon	49
3.3 Μεταβολή Κατώτατου Μισθού και Απασχόληση στο New Jersey	53
3.4 Αύξηση Επιδότησης Εργατικού Ατυχήματος και Διάρκεια Άδειας	62
3.5 Ανακεφαλαίωση	67

4. Οικονομική Κρίση και Αγορά Αυτοκινήτου – Εμπειρική Διερεύνηση	68
4.1 Εισαγωγή	68
4.2 Ιστορική Αναδρομή της Οικονομικής Κρίσης	69
4.3 Παρουσίαση Δεδομένων	74
4.4 Εμπειρική Ανάλυση Δεδομένων	76
4.5 Ανακεφαλαίωση	87
Παραρτήματα	
Παράρτημα Αποτελέσματα κανονικότητας των δειγμάτων κατά Kolmogorov-Smirnov	89
Βιβλιογραφία	98

Κατάλογος Πινάκων

1.1	Πραγματικός ρυθμός αύξησης του Α.Ε.Π, Δ.Α.Α και πληθυσμός 2012	2
1.2	Α.Ε.Π Ελλάδας και Γερμανίας 2004 - 2012	3
1.3	Απόπειρες ανθρωποκτονίας, σύνολο ανέργων και πληθυσμός	4
1.4	Συνολικές πωλήσεις αυτοκινήτων στις Η.Π.Α για τα έτη 2010 - 2013	5
2.1	Επιδράσεις φυσικού πειράματος σε πολλαπλές περιόδους	34
2.2	Επιδράσεις πολιτικής ανά ομάδα και ανά περίοδο	38
3.1	Κύριες επιπτώσεις στην απασχόληση	50
3.2	Μισθοί πριν και μετά τη θεσμοθέτηση κατώτατου μισθού	52
3.3	Σχεδιασμός δείγματος και απώλειες	55
3.4	Κατανομή ταχυφαγείων και μέσοι όροι σημαντικών μεταβλητών	56
3.5	Μέση απασχόληση για κάθε κατάσταση πριν και μετά την αύξηση	59
3.6	Κατανομή δείγματος βάσει τραυματισμού	64
3.7	Εκτίμηση μοντέλου διαφορών σε διαφορές	66
4.1	Ρυθμός ανάπτυξης	70
4.2	Ποσοστό ανεργίας % του εργατικού δυναμικού	71
4.3	Ρυθμός ανάπτυξης Ευρωπαϊκών χωρών	72
4.4	Ετήσιες πωλήσεις αυτοκινήτων	75
4.5	Σύνολο δεδομένων πωλήσεων βενζινοκίνητων οχημάτων	78
4.6	Εκτίμηση συντελεστών του υποδείγματος 4.1	78
4.7	Σύνολο δεδομένων πωλήσεων πετρελαιοκίνητων οχημάτων	79
4.8	Εκτίμηση συντελεστών του υποδείγματος 4.2	80
4.9	Εκτίμηση συντελεστών του υποδείγματος 4.3	81
4.10	Σύνολο δεδομένων μεταβολής πωλήσεων οχημάτων	86
4.11	Εκτίμηση συντελεστών του υποδείγματος 4.4	86

Πανεπιστήμιο Πειραιώς

Κατάλογος Διαγραμμάτων

2.1	Γραφική απεικόνιση της μεθόδου των διαφορών σε διαφορές	42
2.2	Παραβίαση της υπόθεσης των παράλληλων τάσεων	43
3.1	Μισθολογική μεταβολή στο Oregon, γυναίκες, 1913 - 1914	52
3.2	Κατανομή κατώτατου ωρομισθίου πριν τη μεταβολή	57
3.3	Κατανομή κατώτατου ωρομισθίου μετά τη μεταβολή	58
3.4	Υπόθεση ύπαρξης παράλληλης τάσης	60
3.5	Μεταβολή απασχόλησης σε New Jersey και Pennsylvania 1991-1997	61
3.6	Ύψος εβδομαδιαίας αποζημίωσης πριν και μετά τη μεταβολή	63
4.1	Πενταετή CDS, Σεπτέμβριος '09- Φεβρουάριος '11	72
4.2	Μέση τιμή καυσίμων στην Ευρωπαϊκή ήπειρο σε ευρώ ανά λίτρο	73
4.3	Ποσοστιαία % μεταβολή πωλήσεων ανά κατηγορία οχήματος	76

Πανεπιστήμιο Πειραιώς

ΚΕΦΑΛΑΙΟ 1

ΒΑΣΙΚΕΣ ΟΙΚΟΝΟΜΕΤΡΙΚΕΣ ΕΝΝΟΙΕΣ ΚΑΙ ΜΕΘΟΔΟΙ

1.1 Εισαγωγή

Με την ανάπτυξη της επιστήμης της στατιστικής, έχουν δημιουργηθεί νέοι κλάδοι επιστημονικού ενδιαφέροντος οι οποίοι βασίζονται σε αυτήν. Ένας από αυτούς τους κλάδους είναι και αυτός της οικονομετρίας που εστιάζει στην πρόβλεψη και εκτίμηση οικονομικών σχέσεων, στον έλεγχο οικονομικών θεωριών, στην αξιολόγηση και εφαρμογή κυβερνητικών και οικονομικών πολιτικών. Με την ορθή χρήση στατιστικών μεθόδων, προβλέπει μακροοικονομικές πολιτικές όπως τα επιτόκια, το ύψος του πληθωρισμού και το Ακαθάριστο Εγχώριο Προϊόν.

Παρόλο που η οικονομετρία κάνει χρήση μεθόδων των μαθηματικών και της στατιστικής, έχει αναπτυχθεί σαν επιστήμη διαφορετική από αυτές λόγω του τρόπου με τον οποίο γίνεται η συλλογή των δεδομένων. Εστιάζει στην συγκέντρωση και ανάλυση οικονομικών δεδομένων, σε αντίθεση με άλλους κλάδους της στατιστικής, όπου τα δεδομένα λαμβάνονται μέσω ελεγχόμενων πειραμάτων. Από την άλλη μεριά, κοινή συνισταμένη σε αυτούς τους δύο κλάδους αποτελούν η χρήση της απλής και πολλαπλής παλινδρόμησης με ποσοτικές και ποιοτικές μεταβλητές, όπως παρουσιάζεται και στη συνέχεια, με την εστίαση όμως της προσοχής και την ερμηνεία των αποτελεσμάτων, να διαφέρει σημαντικά.

Στο παρόν κεφάλαιο εξετάζονται βασικές οικονομετρικές έννοιες και μέθοδοι, που χρησιμεύουν στην κατανόηση και στην ανάλυση των οικονομικών δεδομένων. Αρχικά παρουσιάζονται οι πιο συχνά εμφανιζόμενες κατηγορίες δεδομένων με τις οποίες ασχολείται η επιστήμη της οικονομετρίας και εν συνεχεία τα κατάλληλα υποδείγματα βάσει των οποίων λαμβάνει χώρα η ανάλυση των δεδομένων αυτών όπως, η ανάλυση παλινδρόμησης, η ανάλυση χρονοσειρών, καθώς και η μεθοδολογία για την ανάλυση δεδομένων πάνελ (*data panels*) μέσω των υποδειγμάτων σταθερών επιδράσεων, πρώτων διαφορών και τυχαίων επιδράσεων.

1.2 Κατηγορίες Δεδομένων

Οι πιο συχνά εμφανιζόμενες κατηγορίες μη-πειραματικών δεδομένων είναι αυτές των διαστρωματικών δεδομένων, των δεδομένων χρονοσειρών, των ομαδοποιημένων διαστρωματικών δεδομένων και των δεδομένων πάνελ.

- **Διαστρωματικά Δεδομένα**

Τα διαστρωματικά δεδομένα αποτελούνται από ένα πλήθος μεταβλητών για κάποιες υπό μελέτη μονάδες, όπως άτομα, επιχειρήσεις, πόλεις, χώρες, σε κάποια συγκεκριμένη χρονική στιγμή ή χρονική περίοδο. Αυτή η μορφή των δεδομένων χρησιμοποιείται ευρύτατα στην Οικονομική και στις Κοινωνικές Επιστήμες και είναι σημαντική για την αξιολόγηση των οικονομικών πολιτικών μέσω στατιστικών ελέγχων υποθέσεων.

Στον Πίνακα 1.1 παρουσιάζεται συνοπτικά η μορφή ενός συνόλου διαστρωματικών δεδομένων που αφορά τον πραγματικό ρυθμό αύξησης του ΑΕΠ για ένα σύνολο χωρών που ανήκουν στη ζώνη του Ευρώ για την χρονική περίοδο του 2012, ο Δείκτης Ανθρώπινης Ανάπτυξης (Δ.Α.Α) όπως εκτιμήθηκε από τον ΟΗΕ το 2013 για το έτος 2012, καθώς και τον πληθυσμό αυτής.

Πίνακας 1.1

Πραγματικός ρυθμός αύξησης ΑΕΠ, Δ.Α.Α και πληθυσμός χωρών της Ε.Ε 2012

A/A	Χώρα	Πραγματικός ρυθμός αύξησης του ΑΕΠ	Δείκτης Ανθρώπινης Ανάπτυξης	Πληθυσμός
1	Βέλγιο	-0.3	0.897	10.666.866
2	Γερμανία	0.7	0.920	82.060.000
3	Εσθονία	3.2	0.846	1.340.602
4	Ελλάδα	-6.4	0.860	11.262.000
5	Ολλανδία	-1.3	0.921	16.855.880
6	Ιρλανδία	0.6	0.916	4.593.100
7	Κύπρος	-2.4	0.848	1.054.400
8	Πορτογαλία	-3.2	0.816	10.676.910
9	Ισπανία	-1.3	0.885	46.609.700

Πηγή: Eurostat και Ο.Η.Ε., 2012

- **Δεδομένα Χρονοσειρών**

Τα Δεδομένα Χρονοσειρών αποτελούνται από παρατηρήσεις που αφορούν μία ή περισσότερες μεταβλητές σε διαχρονική βάση και με συγκεκριμένη συχνότητα. Μερικά παραδείγματα αυτών είναι οι ημερήσιες τιμές των μετοχών, ο μηνιαίος δείκτης τιμών καταναλωτή, το ετήσιο Ακαθάριστο Εγχώριο Προϊόν, τα ετήσια στοιχεία πωλήσεων αυτοκινήτων κ.α. Σε αντίθεση με τη διάταξη των διαστρωματικών δεδομένων, στα δεδομένα χρονοσειρών η χρονική διάταξη των παρατηρήσεων μεταφέρει ενδεχομένως σημαντικές πληροφορίες. Οι πιο συνήθεις συχνότητες εμφάνισης των δεδομένων αυτών είναι η ημερήσια, η εβδομαδιαία, η μηνιαία, η τριμηνιαία και η ετήσια.

Ο αναλυτής που καλείται να επεξεργαστεί μία χρονοσειρά έρχεται αντιμέτωπος με τα συνθετικά στοιχεία αυτής τα οποία και οφείλει να τα απαλείψει έτσι ώστε οι εκτιμήσεις του να είναι ακριβείς. Τα συνθετικά αυτά στοιχεία δεν είναι άλλα από τη διαχρονική τάση, το έντονο εποχικό στοιχείο, την κυκλικότητα κυρίως σε περιόδους μεγαλύτερες του έτους αλλά και την μη-κανονικότητα όπου τα δεδομένα μεταβάλλονται με τρόπο τυχαίο και μη-συστηματικό ο οποίος μάλιστα είναι και απροσδιόριστος. Ένα άλλο χαρακτηριστικό των δεδομένων χρονοσειρών που καθιστά την ανάλυση τους πιο απαιτητική σε σχέση με τα διαστρωματικά δεδομένα, είναι η διαχρονική εξάρτηση των παρατηρήσεων και οι έντονες συσχετίσεις που παρουσιάζονται κυρίως με το πρόσφατο ιστορικό τους.

Στον Πίνακα 1.2 παρουσιάζεται συνοπτικά η μορφή ενός συνόλου δεδομένων χρονοσειρών που αφορά ετήσιες παρατηρήσεις του Ακαθάριστου Εγχώριου Προϊόντος των χωρών Ελλάδας και Γερμανίας, σε δισεκατομμύρια ευρώ για τα έτη 2005 έως και 2013.

Πίνακας 1.2

ΑΕΠ Ελλάδας και Γερμανίας 2003-2012

A/A	Έτος	Ακαθάριστο Εγχώριο Προϊόν Ελλάδας	Ακαθάριστο Εγχώριο Προϊόν Γερμανίας
1	2003	170,9	2147,5
2	2004	183,6	2195,7
3	2005	193,0	2224,4
4	2006	208,6	2313,9
5	2007	223,2	2428,5
6	2008	233,2	2473,8
7	2009	231,1	2374,5
8	2010	222,2	2496,2
9	2011	208,5	2592,6
10	2012	195,0	2645,0

Πηγή: Eurostat, 2003- 2012

- **Ομαδοποιημένα Διαστρωματικά Δεδομένα**

Τα ομαδοποιημένα διαστρωματικά δεδομένα έχουν χαρακτηριστικά τόσο διαστρωματικών δεδομένων όσο και δεδομένων χρονοσειρών. Η δημιουργία τους επιτυγχάνεται με την ένωση διαστρωματικών παρατηρήσεων δύο ή και περισσότερων χρονικών στιγμών ή περιόδων, δίχως όμως οι μετρήσεις να προέρχονται από τις ίδιες μονάδες απαραιτήτως. Αυτό έχει σαν αποτέλεσμα την αύξηση του αριθμού των παρατηρήσεων του υπό μελέτη δείγματος.

Η συγκεκριμένη κατηγορία δεδομένων ενδείκνυται για την ανάλυση και την αξιολόγηση των επιδράσεων από την εφαρμογή μίας νέας κυβερνητικής πολιτικής συλλέγοντας δεδομένα πριν και μετά από την εφαρμογή αυτής. Οι μέθοδοι ανάλυσης που χρησιμοποιούνται για την εν λόγω κατηγορία δεδομένων είναι ίδιες με αυτές των απλών διαστρωματικών παρατηρήσεων, με την διαφορά ότι ο χρόνος λήψης αυτών λαμβάνει μέρος στην ανάλυση ως ξεχωριστή μεταβλητή.

Στον Πίνακα 1.3 παρουσιάζεται συνοπτικά η μορφή ενός συνόλου ομαδοποιημένων διαστρωματικών δεδομένων που αφορά τις απόπειρες ανθρωποκτονίας στην ελληνική επικράτεια για διάφορες περιφέρειες στα έτη 2010 και 2012, ο αριθμός των εγγεγραμμένων αναζητούντων εργασίας ανέργων στα μητρώα του Ο.Α.Ε.Δ., καθώς και ο πληθυσμός της κάθε περιφέρειας.

Πίνακας 1.3

Απόπειρες ανθρωποκτονίας, σύνολο ανέργων και πληθυσμός

<i>Περιφερειακή Ενότητα</i>	<i>Έτος</i>	<i>Απόπειρα Ανθρωποκτονίας</i>	<i>Σύνολο ανέργων αναζητούντων εργασία</i>	<i>Πληθυσμός</i>
Αν. Μακ. Θράκης	2010	8	45.344	608.182
Κεντρικής Μακεδονίας	2010	26	124.941	1.880.058
Δυτικής Μακεδονίας	2010	2	23.394	283.689
Θεσσαλίας	2010	8	40.726	732.762
Ηπείρου	2010	7	20.132	336.856
Ιονίων Νήσων	2010	3	9.371	207.855
Δυτικής Ελλάδας	2012	23	54.640	679.796
Στερεάς Ελλάδας	2012	6	36.712	547.390
Αττικής	2012	60	284.845	3.827.624
Πελοποννήσου	2012	18	30.585	577.903
Νοτίου Αιγαίου	2012	6	13.121	308.975
Βορείου Αιγαίου	2012	1	9.869	199.231
Κρήτης	2012	12	32.608	623.065

Πηγή: Ελληνική Αστυνομία, Ο.Α.Ε.Δ, ΕΛ.ΣΤΑΤ, έτη 2010 και 2012

- **Δεδομένα Πάνελ**

Τα δεδομένα πάνελ έχουν τα ίδια χαρακτηριστικά με αυτά των ομαδοποιημένων διαστρωματικών δεδομένων με την διαφορά ότι η κάθε διαστρωματική μονάδα μελετάται διαχρονικά για δύο ή και περισσότερες περιόδους.

Στον Πίνακα 1.4 παρουσιάζεται συνοπτικά η μορφή ενός συνόλου δεδομένων πάνελ που αφορά τις συνολικές πωλήσεις στην αμερικανική αγορά των μεγαλύτερων αυτοκινητοβιομηχανιών του κόσμου για τα έτη 2010, 2011, 2012 και 2013.

Πίνακας 1.4

Συνολικές πωλήσεις αυτοκινήτων στις Η.Π.Α για τα έτη 2010 - 2013

Κατασκευαστής	2013	2012	2011	2010
Ford	2.403.542	2.160.859	2.057.210	1.755.583
Chevrolet	1.947.125	1.851.646	1.775.812	1.565.632
Toyota	1.923.874	1.764.833	1.396.838	1.488.588
Honda	1.359.876	1.266.569	1.023.986	1.096.874
Nissan	1.131.965	1.021.779	944.073	805.159
Dodge	596.343	524.989	708.650	596.627
Hyundai	720.783	703.007	645.691	538.228
Kia	535.179	557.599	485.492	356.268
Jeep	490.454	474.131	419.349	291.138
GMC	450.901	413.881	397.986	334.981
Volkswagen	407.704	438.133	324.402	256.830
Subaru	424.683	336.441	266.989	263.820
Mercedes	334.350	295.063	261.808	225.007
Mazda	283.947	277.044	250.426	229.566
BMW	309.280	281.460	247.907	220.113
Chrysler	302.492	307.967	221.346	197.446
Lexus	273.847	244.166	198.552	229.329
Buick	205.509	180.408	177.633	155.389
Cadillac	182.543	149.782	152.389	146.925
Acura	165.436	156.216	123.299	133.606
Audi	158.061	139.310	117.561	101.629
Infiniti	116.455	119.877	98.461	103.411
Lincoln	81.694	82.150	85.643	85.828

Πηγή: Αυτοκινητοβιομηχανίες και automotive news data center, 2010-2013

Η διαφορετική φύση των παραπάνω δεδομένων δημιουργεί την ανάγκη ανάπτυξης διαφορετικών μεθόδων για την ανάλυσή τους. Ως εκ τούτου η ανάλυση ενός συνόλου

διαστρωματικών δεδομένων, καθώς και ενός συνόλου ομαδοποιημένων διαστρωματικών δεδομένων λαμβάνει χώρα μέσω της ανάλυσης παλινδρόμησης, η ανάλυση δεδομένων χρονοσειρών μέσω της μεθοδολογίας των Box και Jenkins ενώ και η ανάλυση συνόλων δεδομένων πάνελ με πιο προηγμένες μεθόδους ανάλυσης δεδομένων. Όλες αυτές οι μέθοδοι αναλύονται στη συνέχεια.

1.3 Ανάλυση Παλινδρόμησης

Στην οικονομετρία, όπως και σε άλλους επιστημονικούς κλάδους όπου η στατιστική έχει κυρίαρχο ρόλο, σημαντικό εργαλείο για την ανάλυση των δεδομένων αποτελεί η ανάλυση παλινδρόμησης. Μέσω αυτής, επιδιώκεται από τον ερευνητή η μελέτη ύπαρξης συναρτησιακής σχέσης μεταξύ ενός πλήθους μεταβλητών, με απώτερο σκοπό τη δημιουργία ενός γενικότερου υποδείγματος ικανού να ερμηνεύσει τις όποιες σχέσεις εμφανίζονται, αλλά και να προβλέψει τις επιδράσεις που θα υπάρξουν σε τυχόν μεταβολές στην τιμή μίας μεταβλητής, σε επίπεδο όχι μόνο δείγματος αλλά και πληθυσμού.

Το υπόδειγμα της παλινδρόμησης

Το υπόδειγμα της απλής παλινδρόμησης αποτελεί την πιο απλή εκδοχή της ανάλυσης παλινδρόμησης κατά την οποία διερευνάται η ύπαρξη συναρτησιακής σχέσης μεταξύ δύο μεταβλητών, της εξαρτημένης ή ερμηνευμένης μεταβλητής Y και της ανεξάρτητης ή αλλιώς ερμηνευτικής μεταβλητής X , θεωρώντας ότι ο πληθυσμός ερμηνεύεται από μία γραμμική σχέση της μορφής:

$$Y = \beta_0 + \beta_1 X \quad (1.1)$$

όπου ο συντελεστής β_0 αποτελεί την τεταγμένη της αρχής της εξίσωσης και ο συντελεστής β_1 αποτελεί την κλίση αυτής. Είναι προφανές πως αυτές οι δύο μεταβλητές είναι αδύνατο να συνδέονται με μια αιτιοκρατική σχέση λόγω της φύσης των δεδομένων. Ως εκ τούτου για τις μεταβλητές X και Y με n ζεύγη (x_i, y_i) διαστρωματικών παρατηρήσεων θα ισχύει:

$$y_i = \beta_0 + \beta_1 x_i + u_i \quad (1.2)$$

για $i = 1, 2, \dots, v$, όπου η μεταβλητή u_i ονομάζεται όρος τυχαίου σφάλματος ή διαταρακτικός όρος της σχέσης και αντιπροσωπεύει όλους αυτούς τους παράγοντες, πλην της ανεξάρτητης μεταβλητής X , που επηρεάζουν την εξαρτημένη μεταβλητή Y αλλά δεν έχουν συμπεριληφθεί στο πιο πάνω υπόδειγμα.

Πρωταρχικός στόχος για την δημιουργία ενός αποτελεσματικού προβλεπτικού υποδείγματος της εξαρτημένης μεταβλητής Y μέσω των τιμών που λαμβάνει η ανεξάρτητη μεταβλητή X , είναι η ελαχιστοποίηση των διαφορών $u_i = y_i - (\beta_0 + \beta_1 x_i)$ ώστε τα ζεύγη των διαστρωματικών δεδομένων (x_i, y_i) να κινούνται κοντά στην ευθεία που προκύπτει από την γραφική αναπαράσταση της (1.1) σε ένα ορθοκανονικό σύστημα αξόνων.

Γενικεύοντας το απλό γραμμικό υπόδειγμα και διερευνώντας την ύπαρξη συναρτησιακής σχέσης μεταξύ της εξαρτημένης μεταβλητής Y και πλήθους k ανεξάρτητων μεταβλητών X_1, X_2, \dots, X_k προκύπτει το υπόδειγμα πολλαπλής παλινδρόμησης. Είναι προφανές ότι για $k = 1$ το υπόδειγμα είναι το απλό γραμμικό. Και σε αυτό το σημείο ο πληθυσμός ερμηνεύεται από ένα γραμμικό υπόδειγμα της μορφής:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k \quad (1.3)$$

όπου β_0 είναι η τεταγμένη της αρχής, β_1 η παράμετρος που σχετίζεται με την ανεξάρτητη μεταβλητή X_1 , β_2 η παράμετρος που σχετίζεται με την ανεξάρτητη μεταβλητή X_2 και ούτω καθεξής. Επίσης για κάθε σύνολο τιμών της μορφής $(x_{i1}, x_{i2}, x_{i3}, \dots, x_{ik}, y_i)$ θα ισχύει η σχέση:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + u_i$$

για $i = 1, 2, 3, \dots, v$, όπου η μεταβλητή u_i είναι ο διαταρακτικός όρος που περιέχει το σύνολο των παραγόντων, πλην των ανεξάρτητων μεταβλητών, που επηρεάζουν την εξαρτημένη μεταβλητή αλλά δεν έχουν συμπεριληφθεί στο υπόδειγμα. Η γραφική αναπαράσταση της συνάρτησης (1.3) δεν αφορά ευθεία, αλλά ένα πολυεπίπεδο $k + 1$ διαστάσεων.

Εύρεση βέλτιστου προβλεπτικού υποδείγματος

Η εύρεση του βέλτιστου προβλεπτικού υποδείγματος επιτυγχάνεται με την εκτίμηση των συντελεστών $\beta_0, \beta_1, \dots, \beta_k$ της ευθείας ή του πολυεπίπεδου παλινδρόμησης μέσω της ελαχιστοποίησης των σφαλμάτων. Λόγω του ότι οι δυνατικές προβλέψεις αφορούν

το σύνολο του πληθυσμού η εκτίμηση λαμβάνει χώρα μέσω ενός δείγματος n διαστρωματικών παρατηρήσεων $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$ που αφορούν δείγμα του πληθυσμού αυτού.

Έχουν προταθεί διάφορες μέθοδοι για την ελαχιστοποίηση των σφαλμάτων και συνεπώς την εκτίμηση των συντελεστών $\beta_0, \beta_1, \dots, \beta_k$ της εξίσωσης παλινδρόμησης χωρίς όμως η συναρτησιακή σχέση μεταξύ των μεταβλητών X_1, X_2, \dots, X_k και Y να είναι η βέλτιστη. Τελικώς η ορθότερη μέθοδος η οποία και έχει επικρατήσει είναι αυτή των ελάχιστων τετραγώνων μέσω της οποίας επιδιώκεται η ελαχιστοποίηση της ποσότητας του αθροίσματος των τετραγώνων των σφαλμάτων όπως φαίνεται και στην παρακάτω σχέση:

$$\varphi(\beta_0, \beta_1, \beta_2, \dots, \beta_k) = \sum_{i=1}^n \hat{u}_i^2 = \sum_{i=1}^n [y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \dots + \hat{\beta}_k x_{ik})]^2$$

Αυτό επιτυγχάνεται με την εξίσωση των μερικών παραγώγων για κάθε ένα από τα β_j με $j = 0, 1, 2, \dots, k$ με το μηδέν. Μετά από αυτή τη διαδικασία οδηγούμαστε σε $k + 1$ γραμμικές εξισώσεις με $k + 1$ άγνωστες παραμέτρους. Οι εξισώσεις αυτές είναι της μορφής:

$$\begin{aligned} \sum_{i=1}^n (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_k x_{ik}) &= 0 \\ \sum_{i=1}^n x_{i1} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_k x_{ik}) &= 0 \\ \sum_{i=1}^n x_{i2} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_k x_{ik}) &= 0 \\ &\vdots \\ \sum_{i=1}^n x_{ik} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_{i1} - \dots - \hat{\beta}_k x_{ik}) &= 0 \end{aligned}$$

Για την επίλυση του $(k + 1) \times (k + 1)$ συστήματος δημιουργούμε τον πίνακα σχεδιασμού X

του υποδείγματος που περιλαμβάνει τις τιμές των ανεξάρτητων μεταβλητών και τα διανύσματα \mathbf{y} και $\boldsymbol{\beta}$. Ο πίνακας X είναι διάστασης $n \times (k + 1)$, ο \mathbf{y} διάστασης $n \times 1$ και ο $\boldsymbol{\beta}$ διάστασης $(k + 1) \times 1$.

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ 1 & x_{31} & x_{32} & \cdots & x_{3k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{v1} & x_{v2} & \cdots & x_{vk} \end{bmatrix}, \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_v \end{bmatrix}, \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$$

Με κατάλληλους μετασχηματισμούς προκύπτει τελικά ότι το παραπάνω σύστημα μπορεί να γραφεί με μορφή πινάκων ως $(X'X)\hat{\boldsymbol{\beta}} = X'y$ και πολλαπλασιάζοντας και τα δύο μέλη με τον πίνακα $(X'X)^{-1}$ προκύπτουν τελικά οι εκτιμήτριες ελαχίστων

τετραγώνων $\hat{\boldsymbol{\beta}} = (X'X)^{-1}X'y$, όπου $\hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \hat{\beta}_3 \\ \vdots \\ \hat{\beta}_k \end{bmatrix}$.

Το υπόδειγμα της πολλαπλής παλινδρόμησης με τη χρήση πινάκων είναι της μορφής $\mathbf{y} = X\hat{\boldsymbol{\beta}} + \mathbf{u}$, όπου το \mathbf{u} με τη σειρά του είναι το διάνυσμα των διαταρακτικών όρων,

$\mathbf{u} = \begin{bmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_k \end{bmatrix}$ με διάσταση $(k+1) \times 1$. Τελικώς η εξίσωση της συνάρτησης

παλινδρόμησης έχει τη μορφή:

$$\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \cdots + \hat{\beta}_k x_{ik}$$

για $i = 1, 2, 3, \dots, v$.

Όπως αναφέρθηκε και πιο πάνω οι παρατηρήσεις $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$ είναι ανέφικτο να βρίσκονται ακριβώς πάνω στο πολυεπίπεδο παλινδρόμησης λόγω του ότι η εξαρτημένη μεταβλητή Y επηρεάζεται από διάφορους παράγοντες και όχι μόνον από τις ανεξάρτητες μεταβλητές X_j , για $j = 1, \dots, k$. Η απόκλιση των εκτιμηθέντων τιμών της εξαρτημένης μεταβλητής Y από τις παρατηρούμενες τιμές της, αποτελούν τα κατάλοιπα του υποδείγματος τα οποία είναι οι εκτιμώμενες τιμές των σφαλμάτων $\hat{u}_i = y_i - \hat{y}_i$.

Υποθέσεις του γραμμικού υποδείγματος

Σύμφωνα με το θεώρημα των Gauss - Markov οι εκτιμητές των ελαχίστων τετραγώνων είναι BLUE (*Best Linear Unbiased Estimators*), έχουν δηλαδή τη μικρότερη διακύμανση μεταξύ του πλήθους των αμερόληπτων εκτιμητών που είναι γραμμικές συναρτήσεις των Y_i . Για την εξασφάλιση της αμεροληψίας και κατ' επέκταση την ορθή εκτίμηση των παραμέτρων είναι υποχρεωτική η τήρηση μίας σειράς υποθέσεων η οποία και επηρεάζει άμεσα την αξιοπιστία των αποτελεσμάτων της συνάρτησης παλινδρόμησης. Υποθέτοντας ότι το υπόδειγμα είναι της μορφής:

$$Y = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k + u$$

οι υποθέσεις είναι:

1^η υπόθεση: Η αναμενόμενη τιμή των σφαλμάτων u_i είναι ίση με 0. Συνεπώς, $E(u_i | x_1, \dots, x_k) = 0$.

2^η υπόθεση: Η διακύμανση των σφαλμάτων είναι σταθερή και ίση με σ^2 . Επομένως, $V(u_i | x_1, \dots, x_k) = \sigma^2$. Η υπόθεση αυτή ονομάζεται και «ομοσκεδαστικότητα».

3^η υπόθεση: Τα σφάλματα δε σχετίζονται με τις ανεξάρτητες μεταβλητές. Επομένως, $Cov(u_i, \mathbf{X}) = 0$, όπου \mathbf{X} το διάνυσμα των ανεξάρτητων μεταβλητών.

4^η υπόθεση: Τα σφάλματα είναι ανεξάρτητα μεταξύ τους.

5^η υπόθεση: Από τις ανεξάρτητες μεταβλητές X_1, X_2, \dots, X_k καμία δεν είναι σταθερή και μεταξύ τους είναι ανεξάρτητες έτσι ώστε να μην υπάρχει το πρόβλημα της πολυσυγγραμμικότητας.

6^η υπόθεση: Τα σφάλματα κατανέμονται κανονικά με τη μέση τιμή και τη διακύμανση που προαναφέρθηκε. Επομένως, $u_i \sim iidN(0, \sigma^2)$.

Η ισχύς της υπόθεσης της κανονικότητας είναι πολύ σημαντική. Η κατανομή των εκτιμητών ελαχίστων τετραγώνων εξαρτάται από την κατανομή των σφαλμάτων και έτσι κατανέμονται και αυτοί κανονικά με μέση τιμή β_j και διακύμανση $Var(\hat{\beta}_j) = \sigma^2 (X'X)^{-1}$. Λόγω της κανονικής κατανομής ο εκάστοτε ερευνητής δύναται να δημιουργήσει μέσω απλών διαδικασιών διαστήματα εμπιστοσύνης για τις παραμέτρους $\beta_j, j = 0, \dots, k$, αλλά και ελέγχους υποθέσεων με μεγάλο πρακτικό ενδιαφέρον. Το υπόδειγμα που ικανοποιεί το σύνολο των παραπάνω υποθέσεων ονομάζεται κανονικό γραμμικό υπόδειγμα, είναι της μορφής:

$$E(Y | X_1, \dots, X_k) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$$

και

$$V(Y|X_1, \dots, X_k) = \sigma^2$$

και περιγράφει την κατά μέσο όρο συμπεριφορά της εξαρτημένης μεταβλητής συναρτήσει των ανεξάρτητων μεταβλητών, καθώς και τυχόν αποκλίσεις που υπάρχουν από το υπόδειγμα.

Ποιότητα προσαρμογής

Μετά τη δημιουργία υποδείγματος παλινδρόμησης είναι χρήσιμη η ποσοτικοποίηση της προσαρμογής της ευθείας ή του πολυεπίπεδου παλινδρόμησης στα δεδομένα. Ορίζοντας ως $SST = \sum_{i=1}^n (y_i - \bar{y})^2$ το συνολικό άθροισμα τετραγώνων που εκφράζει τη μεταβλητότητα των παρατηρούμενων τιμών y_i ως προς τη μέση τους τιμή \bar{y} , $SSE = \sum_{i=1}^n \hat{u}_i^2$ το άθροισμα τετραγώνων των σφαλμάτων που εκφράζει την απόσταση των παρατηρούμενων τιμών y_i από τις εκτιμώμενες \hat{y}_i και ως $SSR = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2$ το άθροισμα τετραγώνων της παλινδρόμησης που εκφράζει την απόσταση των εκτιμώμενων τιμών \hat{y}_i από τη μέση τιμή \bar{y} , αποδεικνύεται ότι $SST = SSR + SSE$ και επομένως, προκύπτει ο συντελεστής προσδιορισμού $R^2 = \frac{SSR}{SST}$ ο οποίος δείχνει ποσοτικά την ποιότητα της προσαρμογής στα δεδομένα. Λόγω της σχέσης $0 \leq SSR \leq SST$, ο συντελεστής προσδιορισμού λαμβάνει τιμές στο διάστημα $[0,1]$. Πιο συγκεκριμένα $R^2 = 0$ όταν η ευθεία ή το πολυεπίπεδο παλινδρόμησης δεν έχει καμία προβλεπτική ικανότητα ενώ $R^2 = 1$ όταν περιγράφει τέλεια τα δεδομένα.

Για το πολλαπλό γραμμικό υπόδειγμα χρησιμοποιείται κυρίως ο τροποποιημένος συντελεστής προσδιορισμού ο οποίος λαμβάνει υπόψη και το πλήθος των $k + 1$ παραμέτρων του υποδείγματος όπως φαίνεται στη συνέχεια:

$$R_{adj}^2 = 1 - \left[\frac{SSE}{v - k - 1} / \frac{SST}{v - 1} \right]$$

Επομένως, για τη δημιουργία ενός αξιόπιστου προβλεπτικού υποδείγματος απαιτείται ο συντελεστής προσδιορισμού να τείνει στη μονάδα. Στις κοινωνικές επιστήμες και στην οικονομετρία παρόλα αυτά η ύπαρξη χαμηλού συντελεστή προσδιορισμού είναι σύνηθες φαινόμενο. Κάτι τέτοιο δε σημαίνει υποχρεωτικά ότι η συνάρτηση παλινδρόμησης δεν περιγράφει με ικανοποιητικό τρόπο τα δεδομένα της

ανάλυσης. Είναι πολύ πιθανό οι ερμηνευτικές μεταβλητές να ερμηνεύουν ένα σημαντικό μέρος της διακύμανσης του δείγματος στην εξαρτημένη μεταβλητή Y .

Σημαντικότητα εκτιμητών ελαχίστων τετραγώνων

Στο κανονικό γραμμικό υπόδειγμα μελετάται η αναμενόμενη τιμή της εξαρτημένης μεταβλητής συναρτήσει των τιμών των ανεξάρτητων μεταβλητών. Πιο συγκεκριμένα το υπόδειγμα του πληθυσμού έχει τη μορφή $E(Y) = \beta_0 + \beta_1 X_1 + \dots + \beta_k X_k$ όπου κάθε μία από τις παραμέτρους β_j εκφράζει την αναμενόμενη μεταβολή της Y αν η ανεξάρτητη μεταβλητή X_j αυξηθεί κατά μία μονάδα και οι υπόλοιπες $X_1, X_2, \dots, X_{j-1}, X_{j+1}, \dots, X_k$ παραμείνουν σταθερές. Λόγω του ότι οι παράμετροι β_j αποτελούν άγνωστα χαρακτηριστικά του πληθυσμού που εκτιμώνται σημειακά μέσω της μεθόδου ελαχίστων τετραγώνων, η δημιουργία ελέγχου υποθέσεων για κάποια από τις παραμέτρους, αποτελεί πολύ σημαντικό κομμάτι της ανάλυσης. Συγκεκριμένα το ενδιαφέρον του ερευνητή εστιάζεται στον έλεγχο της υπόθεσης $H_0: \beta_j = 0$ έναντι μίας αμφίπλευρης ή μονόπλευρης εναλλακτικής. Ο παραπάνω έλεγχος μπορεί να πραγματοποιηθεί είτε μέσω του t test, είτε μέσω της τιμής p , είτε με τη δημιουργία διαστήματος εμπιστοσύνης.

1.4 Χρονοσειρές

Η ανάλυση χρονοσειρών (*time series analysis*) είναι το πεδίο έρευνας που μελετά διαδικασίες που εξελίσσονται διαχρονικά. Σκοπός της εν λόγω ανάλυσης είναι η μελέτη και αναγνώριση της φύσης ενός φαινομένου αλλά και η πρόβλεψη της μελλοντικής εξέλιξης του φαινομένου με τη χρήση παρελθοντικών τιμών των ίδιων των μεταβλητών και πιθανόν παρούσες και παρελθοντικές τιμές ενός όρου σφάλματος. Στοιχεία που διαδραματίζουν σημαντικότατο ρόλο στην ανάλυση χρονοσειρών είναι η τάση (*trend*), η εποχικότητα (*seasonality*), οι αυτοσυσχετίσεις (*autocorrelations*) και τα κατάλοιπα ή ο διαταρακτικός όρος u_t .

Σε αντίθεση με τα διαστρωματικά δεδομένα όπου συλλέγονται κατά την ίδια χρονική περίοδο, τα δεδομένα μίας χρονοσειράς λαμβάνονται σε ισαπέχοντα χρονικά σημεία ή χρονικές περιόδους. Πιο συγκεκριμένα μία χρονοσειρά είναι ένα δείγμα y_1, y_2, \dots, y_T παρατηρήσεων, όπου ο δείκτης T παριστάνει τις ισαπέχουσες χρονικές

περιόδους και η κάθε τιμή του δείγματος αντιστοιχεί σε συγκεκριμένη τιμή μίας τυχαίας μεταβλητής Y_1, Y_2, \dots, Y_T . Το πλήθος των τυχαίων μεταβλητών είναι μέρος μίας άπειρης ακολουθίας τυχαίων μεταβλητών η οποία ονομάζεται στοχαστική διαδικασία και συμβολίζεται ως $\{y_T\}$. Στα υποδείγματα χρονοσειρών η τρέχουσα τιμή μιας οικονομικής τυχαίας μεταβλητής Y εκφράζεται ως συνάρτηση των προηγούμενων τιμών της, δηλαδή των τιμών της με χρονική υστέρηση, ενώ σε ένα υπόδειγμα παλινδρόμησης η τυχαία μεταβλητή Y είναι συνάρτηση k ερμηνευτικών μεταβλητών.

Στασιμότητα Χρονοσειρών

Οι χρονοσειρές διακρίνονται σε στάσιμες και μη στάσιμες. Η ιδιότητα της στασιμότητας είναι ίσως η σημαντικότερη μιας και έχει πρωτεύοντα ρόλο στη δημιουργία ενός ορθού μαθηματικού υποδείγματος. Διαισθητικά μία χρονοσειρά θεωρείται στάσιμη εάν δεν υπάρχει συστηματική αλλαγή του μέσου όρου και της διασποράς της στο χρόνο. Στη συνέχεια αναφέρεται η μαθηματική προσέγγιση των στάσιμων διαδικασιών, καθώς και οι συνήθεις στάσιμες διαδικασίες του κινητού μέσου (MA), η αυτοπαλίνδρομη διαδικασία (AR), η μικτή διαδικασία (ARMA) και η αυτοπαλίνδρομη ολοκληρωμένη διαδικασία κινητού μέσου ARIMA κατά την οποία μία μη στάσιμη χρονοσειρά μετασχηματίζεται σε στάσιμη σύμφωνα με τη μεθοδολογία των Box και Jenkins (1976).

Αυστηρώς Στάσιμη Διαδικασία

Αυστηρώς στάσιμη (*Strictly Stationary Process*) ονομάζεται η διαδικασία για την οποία για κάθε $t_1, t_2, \dots, t_T \in \mathbb{Z}$ κάθε $k \in \mathbb{Z}$ και $T \in 1, 2, \dots$ ισχύει:

$$F_{y_{t_1}, y_{t_2}, \dots, y_{t_T}}(y_1, \dots, y_T) = F_{y_{t_1+k}, y_{t_2+k}, \dots, y_{t_T+k}}(y_1, \dots, y_T)$$

όπου η F δηλώνει την από κοινού συνάρτηση κατανομής του συνόλου των τυχαίων μεταβλητών (Tong, 1990). Μπορεί επίσης να αναφερθεί ότι το μέτρο της πιθανότητας για την ακολουθία $\{y_t\}$ είναι το ίδιο με αυτό της ακολουθίας $\{y_{t+k}\} \forall k$, όπου το k δηλώνει μία αυθαίρετη μετακίνηση κατά μήκος του άξονα του χρόνου είτε προς το μέλλον είτε προς το παρελθόν. Με άλλα λόγια, μία διαδικασία είναι αυστηρώς στάσιμη εάν η κατανομή των τιμών της παραμένει ίδια και κατ' επέκταση και οι ιδιότητές της, καθώς ο χρόνος μεταβάλλεται.

Ο αυστηρός ορισμός της στασιμότητας αναφέρεται σε όλες τις ιδιότητες μιας στοχαστικής διαδικασίας, γεγονός που καθιστά την εύρεση στασιμότητας αρκετά δύσκολη διαδικασία. Για το λόγο αυτό για την ανάλυση μίας χρονοσειράς αρκεί να πληρούνται κάποιες μόνο προϋποθέσεις για να είναι η διαδικασία ασθενώς στάσιμη, που σχετίζονται με η μέση τιμή, τη διακύμανση και τις συνδιακυμάνσεις των παρατηρήσεων της χρονοσειράς, όπως ορίζεται στη συνέχεια.

Ασθενώς Στάσιμη Διαδικασία

Ασθενώς στάσιμη (*Strictly Stationary Process*) ονομάζεται η διαδικασία για την οποία ικανοποιούνται οι παρακάτω συνθήκες για $t = 1, 2, \dots, \infty$:

$$(1) E(Y_t) = \mu, \text{ ανεξάρτητη του } t$$

$$(2) Var(Y_t) = \sigma^2 < \infty, \text{ ανεξάρτητη του } t$$

$$(3) Cov(Y_t, Y_{t+k}) = Cov(Y_{t+m}, Y_{t+m+k}) = \gamma_k \quad \forall k, \text{ ανεξάρτητη του } t$$

Αυτές οι τρεις εξισώσεις δηλώνουν ότι μία στάσιμη διαδικασία έχει διαχρονικά σταθερό μέσο και σταθερή διακύμανση και οι συνδιακυμάνσεις της μεταξύ των y_t και y_{t+k} είναι συναρτήσεις μόνο της χρονικής υστέρησης k .

Από τη σχέση (3) είναι προφανές ότι $Cov(Y_t, Y_{t+k}) = Cov(Y_{t-k}, Y_t)$ δηλαδή $\gamma_k = \gamma_{-k}$ και για $k = 0$, $\gamma_0 = Cov(Y_t, Y_t) = Var(Y_t) = \sigma^2$. Επειδή η συνδιακύμανση αφορά παρατηρήσεις της ίδιας μεταβλητής που απέχουν χρονικά μεταξύ τους κατά k , η συνδιακύμανση ονομάζεται και αυτοσυνδιακύμανση.

Η μέση τιμή, η διακύμανση και η αυτοσυνδιακύμανση αποτελούν άγνωστες προς εκτίμηση παραμέτρους. Επομένως, για δείγμα T περιόδων προκύπτει:

- $\hat{\mu} = \frac{1}{T} \sum_{t=1}^T Y_t$
- $\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\mu})^2$
- $\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\mu})(Y_{t+k} - \hat{\mu})$

Οι αυτοσυνδιακυμάνσεις δεν είναι ιδιαίτερα χρήσιμες για τη μέτρηση της σχέσης της Y με τις παρελθοντικές τιμές της μιας και εξαρτάται από τις μονάδες μέτρησης της y_t με αποτέλεσμα οι τιμές που προκύπτουν να μην έχουν καμία άμεση ερμηνεία. Για το λόγο αυτό, είναι πιο βολικό να χρησιμοποιούνται οι αυτοσυσχετίσεις (*autocorrelations*) που αποτελούν την κανονικοποίηση των αυτοσυνδιακυμάνσεων. Πιο συγκεκριμένα ο συντελεστής αυτοσυσχέτισης δύο μεταβλητών ορίζεται ως ο λόγος

της συνδιακύμανσης προς το γινόμενο των τετραγωνικών ριζών των διακυμάνσεων των δύο αυτών μεταβλητών. Έτσι για τις μεταβλητές Y_t και Y_{t+k} προκύπτει η σχέση:

$$\rho_k = \frac{Cov(Y_t, Y_{t+k})}{\sqrt{Var(Y_t)} \cdot \sqrt{Var(Y_{t+k})}} = \frac{\gamma_k}{\gamma_0}$$

η οποία δεν εξαρτάται από το χρόνο t αλλά μόνο από την υστέρηση k και λαμβάνει τιμές στο κλειστό διάστημα $[-1,1]$. Η σχέση που υπάρχει ανάμεσα στο συντελεστή αυτοσυσχέτισης ρ και στη χρονική υστέρηση k ονομάζεται συνάρτηση αυτοσυσχέτισης (*Autocorrelation Function ACF*) και η γραφική της απεικόνιση διάγραμμα αυτοσυσχέτισης (*correlogram*), ενώ η εκτίμηση του συντελεστή αυτοσυσχέτισης δίνεται από τη σχέση:

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^T (Y_t - \hat{\mu})(Y_{t+k} - \hat{\mu})}{\sum_{t=1}^T (Y_t - \hat{\mu})^2}$$

Στην ανάλυση χρονοσειρών η σημασία της συνάρτησης αυτοσυσχέτισης είναι πολύ μεγάλη, γιατί δείχνει τόσο το βαθμό όσο και το μήκος ή τη χρονική διάρκεια της μνήμης της στοχαστικής διαδικασίας

Διαδικασία Λευκού Θορύβου

Μία διαδικασία λευκού θορύβου (*white noise process*) δεν έχει σαφή δομή και εμφανίζεται όταν ισχύουν τα παρακάτω:

- (1) $E(Y_t) = \mu$
- (2) $Var(Y_t) = \sigma^2$
- (3) $Cov(Y_t, Y_s) = 0$, για $t \neq s$

Επομένως, η διαδικασία λευκού θορύβου, έχει σταθερή μέση τιμή και διακύμανση, ενώ οι αυτοσυνδιακυμάνσεις της είναι μηδενικές. Στην πράξη σε μία διαδικασία λευκού θορύβου κάθε παρατήρηση είναι ασυσχέτιστη με όλες τις υπόλοιπες τιμές της ακολουθίας.

Αυτοπαλίνδρομες διαδικασίες

Αυτοπαλίνδρομη (*autoregressive process*) ονομάζεται η διαδικασία της οποίας η παρούσα τιμή μίας μεταβλητής y , εξαρτάται μόνο από τις παρελθοντικές τιμές της

συμπεριλαμβανομένου και ενός όρου σφάλματος. Ένα αυτοπαλίνδρομο υπόδειγμα τάξης p , συμβολίζεται με AR(p) και εκφράζεται με τη σχέση:

$$y_t = \mu + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + u_t \quad (1.4)$$

όπου ο όρος u_t αναφέρεται στο θόρυβο. Εισάγοντας στο υπόδειγμα τον τελεστή χρονικής υστέρησης L (*Lag Operator*) η σχέση (1.4) γίνεται:

$$y_t = \mu + \sum_{i=1}^p \varphi_i L^i y_t + u_t$$

ή εναλλακτικά

$$\varphi(L)y_t = \mu + u_t \quad (1.5)$$

όπου $L^i y_t = y_{t-i}$ και $\varphi(L) = (1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p)$.

Η στασιμότητα αποτελεί μία επιθυμητή ιδιότητα σε ένα εκτιμώμενο αυτοπαλίνδρομο υπόδειγμα, για πολλούς λόγους. Ένας σημαντικός λόγος είναι ότι σε ένα υπόδειγμα με μη στάσιμους συντελεστές, οι προγενέστερες τιμές του όρου σφάλματος, έχουν μη πτωτική επίδραση στην παρούσα τιμή της μεταβλητής y_t καθώς ο χρόνος προχωράει, πράγμα απίθανο σε πολλές εφαρμογές.

Θέτοντας για χάρη ευκολίας τη μέση τιμή μ ίση με το μηδέν στην (1.5), το υπόδειγμα ονομάζεται αυτοπαλίνδρομο υπόδειγμα q τάξης μηδενικού μέσου και περιγράφεται από τη σχέση:

$$\varphi(L)y_t = u_t \quad (1.6)$$

Πρέπει να αναφερθεί ότι η διαδικασία (1.6) θα είναι στάσιμη αν μπορεί να γραφεί στη μορφή:

$$y_t = \varphi(L)^{-1} u_t$$

με τη συνάρτηση $\varphi(L)^{-1}$ να συγκλίνει στο μηδέν. Αυτό έχει σαν αποτέλεσμα οι αυτοσυσχετίσεις να μειώνονται καθώς η χρονική υστέρηση αυξάνεται. Όταν υπολογίζεται η $\varphi(L)^{-1}$, περιλαμβάνει άπειρο πλήθος όρων και μπορεί να γραφεί και σαν MA(∞). Αν η διαδικασία (1.3.3) είναι στάσιμη, οι συντελεστές στην MA(∞) θα μειώνονται κατά μήκος της χρονικής υστέρησης, ενώ αν η διαδικασία είναι μη στάσιμη, οι συντελεστές δεν θα συγκλίνουν στο μηδέν καθώς το μήκος της χρονικής υστέρησης αυξάνεται.

Η συνθήκη για τον έλεγχο της ύπαρξης στασιμότητας σε ένα αυτοπαλίνδρομο υπόδειγμα p τάξης είναι οι ρίζες της χαρακτηριστικής συνάρτησης του πολυωνύμου:

$$1 - \varphi_1 z - \varphi_2 z^2 - \dots - \varphi_p z^p = 0$$

να βρίσκονται εκτός του μοναδιαίου κύκλου.

Διαδικασία Κινητού Μέσου (Moving Average Process)

Η απλούστερη κατηγορία υποδειγμάτων χρονοσειρών είναι αυτή του κινητού μέσου. Υποθέτοντας ότι ο διαταρακτικός όρος u_t ($t = 1, 2, 3, \dots$) αποτελεί μία διαδικασία λευκού θορύβου με $E(u_t) = 0$ και $Var(u_t) = \sigma^2$ δημιουργείται το υπόδειγμα:

$$y_t = \mu + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q} \quad (1.7)$$

το οποίο ονομάζεται υπόδειγμα κινητού μέσου q τάξης και συμβολίζεται με $MA(q)$. Το παρόν υπόδειγμα είναι στάσιμο μιας και αποτελείται από πεπερασμένο αριθμό όρων λευκού θορύβου και εκφράζει τη βραχυχρόνια συμπεριφορά της χρονοσειράς. Εισάγοντας στο υπόδειγμα τον τελεστή χρονικής υστέρησης, εξετάζεται η μακροχρόνια συμπεριφορά της σειράς και η σχέση (1.7) μπορεί να γραφεί ως:

$$y_t = \mu + u_t - \theta_1 L^1 u_t - \theta_2 L^2 u_t - \dots - \theta_q L^q u_t$$

ή εναλλακτικά ως:

$$y_t = \mu + \theta(L)u_t$$

όπου: $\theta(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \dots - \theta_q L^q$.

Στις περισσότερες εφαρμογές ο σταθερός όρος μ δεν συμπεριλαμβάνεται στις εξισώσεις. Απαλείφοντάς τον διευκολύνεται σημαντικά η πολυπλοκότητα της άλγεβρας χωρίς να υπάρχει απώλεια της γενικότητας. Υποθέτοντας ένα δείγμα παρατηρήσεων μίας διαδικασίας $\{z_t\}$ με μέση τιμή \bar{z} , η δημιουργία ενός υποδείγματος κινητού μέσου χωρίς το σταθερό όρο προκύπτει από την αφαίρεση της μέσης τιμής \bar{z} από κάθε παρατήρηση z_t .

Οι χαρακτηριστικές ιδιότητες μίας διαδικασίας κινητού μέσου τάξης q αναφέρονται στη συνέχεια:

$$(1) E(Y_t) = \mu$$

$$(2) Var(Y_t) = \gamma_0 = (1 + \theta_1^2 + \theta_2^2 + \dots + \theta_q^2)\sigma^2$$

$$(3) \gamma_s = \begin{cases} (\theta_s + \theta_{s+1}\theta_1 + \theta_{s+2}\theta_2 + \dots + \theta_q\theta_{q-s})\sigma^2, & s = 0, 1, 2, \dots, q \\ 0, & s > q \end{cases}$$

Επομένως, μία διαδικασία κινητού μέσου έχει σταθερή μέση τιμή και διακύμανση και μη μηδενικές αυτοσυνδιακυμάνσεις για χρονική υστέρηση ίση με q , ενώ μηδενικές αυτοσυνδιακυμάνσεις σε κάθε άλλη περίπτωση.

Σε ένα υπόδειγμα κινητού μέσου q τάξης $MA(q)$, οι ρίζες της χαρακτηριστικής εξίσωσης $\theta(z) = 0$ πρέπει να είναι μεγαλύτερες της μονάδας κατά απόλυτη τιμή. Η

κατάσταση αυτή ονομάζεται αντιστρεψιμότητα και ερμηνεύεται με τον ίδιο τρόπο με αυτήν της στασιμότητας με τη διαφορά ότι η πρώτη αναφέρεται σε υποδείγματα κινητού μέσου, ενώ η δεύτερη σε αυτοπαλίνδρομα υποδείγματα. Επομένως, όταν έχει εξασφαλιστεί η αντιστρεψιμότητα η διαδικασία κινητού μέσου q τάξης μπορεί να παρουσιαστεί μακροχρόνια από μια αυτοπαλίνδρομη διαδικασία απείρων όρων $AR(\infty)$.

Διαδικασίες ARMA

Με το συνδυασμό ενός αυτοπαλίνδρομου υποδείματος p τάξης $AR(p)$ με ένα υπόδειγμα κινητού μέσου q τάξης $MA(q)$ προκύπτει ένα υπόδειγμα $ARMA(p,q)$ στο οποίο η παρούσα τιμή μίας διαδικασίας y εξαρτάται γραμμικά από τις παρελθοντικές τιμές της και από έναν συνδυασμό παρούσων και παρελθοντικών τιμών του λευκού θορύβου. Το υπόδειγμα μπορεί να γραφεί στη μορφή:

$$\varphi(L)y_t = \mu + \theta(L)u_t \quad (1.8)$$

όπου

$$\varphi(L) = (1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) \text{ και}$$

$$\theta(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \dots - \theta_q L^q$$

ή

$$y_t = \mu + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \dots - \theta_q u_{t-q} + u_t$$

με

$$E(u_t) = 0, E(u_t^2) = \sigma^2, E(u_t u_s) = 0, t \neq s$$

Τα χαρακτηριστικά μίας διαδικασίας ARMA είναι ένας συνδυασμός των χαρακτηριστικών της αυτοπαλίνδρομης διαδικασίας και αυτής του κινητού μέσου. Σε αυτό το σημείο σημαντικό ρόλο έχει τόσο η συνάρτηση αυτοσυσχέτισης (*ACF*) όσο και η συνάρτηση μερικής αυτοσυσχέτισης (*Partial Autocorrelation Function PACF*) με τη δεύτερη να αναφέρεται στη συσχέτιση της παρούσας τιμής της y_t με την αντίστοιχη τιμή της k περιόδους πίσω, αφαιρώντας τις γραμμικές επιδράσεις των ενδιάμεσων τιμών $y_{t-1}, y_{t-2}, \dots, y_{t-(k-1)}$ και βάσει της οποίας καθορίζεται η τάξη του καλύτερα προσαρμοζόμενου υποδείματος.

Συγκεντρώνοντας τα χαρακτηριστικά των τριών διαδικασιών που έχουν αναφερθεί ως τώρα προκύπτουν τα παρακάτω:

Αυτοπαλίνδρομη Διαδικασία:

- Η συνάρτηση αυτοσυσχέτισης φθίνει γεωμετρικά με την πάροδο του χρόνου
- Ένας αριθμός μη μηδενικών σημείων της συνάρτησης μερικής αυτοσυσχέτισης = τάξη της αυτοπαλίνδρομης διαδικασίας

Διαδικασία Κινητού Μέσου:

- Ένας αριθμός μη μηδενικών σημείων της συνάρτησης αυτοσυσχέτισης = τάξη της διαδικασίας κινητού μέσου
- Η συνάρτηση μερικής αυτοσυσχέτισης φθίνει γεωμετρικά με την πάροδο του χρόνου

Διαδικασία ARMA:

- Η συνάρτηση αυτοσυσχέτισης φθίνει γεωμετρικά με την πάροδο του χρόνου
- Η συνάρτηση μερικής αυτοσυσχέτισης φθίνει γεωμετρικά με την πάροδο του χρόνου

Διαδικασίες ARIMA

Τα τρία στοχαστικά υποδείγματα που εξετάστηκαν έως τώρα αναφέρονται όλα σε στάσιμες διαδικασίες που σημαίνει ότι ο μέσος και η διακύμανση παραμένουν σταθεροί, ενώ οι αυτοσυνδιακυμάνσεις εξαρτώνται μόνο από τη χρονική υστέρηση. Δυστυχώς, οι περισσότερες αν όχι όλες οι οικονομικές χρονοσειρές, όπως ο δείκτης κατανάλωσης, η ανεργία, ο δείκτης τιμών, τα κέρδη δεν έχουν τα χαρακτηριστικά στάσιμων διαδικασιών. Τη λύση σε αυτό το πρόβλημα την έφεραν οι Box και Jenkins οι οποίοι πρότειναν το μετασχηματισμό της μη στάσιμης χρονοσειράς, λαμβάνοντας απλά τις διαφορές της κάθε τιμής με την τιμή της προηγούμενης περιόδου τόσες φορές έως ότου να γίνει στάσιμη. Αποτέλεσμα αυτής της διαδικασίας είναι μία ολοκληρωμένη d τάξης στάσιμη χρονοσειρά (*integrated d^{th} order*) η οποία δημιουργείται μετά από πλήθος d διαφορών και συμβολίζεται με $I(d)$.

Ένα υπόδειγμα ARMA(p, q) που εφαρμόζεται σε μία ολοκληρωμένη σειρά d τάξεως, ονομάζεται αυτοπαλίνδρομο ολοκληρωμένο υπόδειγμα κινητού μέσου τάξεως (p, d, q) (Autoregressive Integrated Moving Average) και συμβολίζεται με ARIMA(p, d, q). Συνεπώς, μία διαδικασία ARIMA(p, d, q) είναι μία διαδικασία η οποία διαφορίζεται d φορές και παράγει μία ARMA(p, q).

Έστω ότι για μία μη στάσιμη χρονοσειρά y_t απαιτείται d πλήθος διαφορών έτσι ώστε να γίνει στάσιμη. Με τη χρήση του τελεστή υστέρησης αποδεικνύεται ότι:

$$\Delta^d y_t = (1 - L)^d y_t \quad (1.9)$$

Με χρήση της σχέσης (1.8) αλλά και της (1.9) το μαθηματικό υπόδειγμα μίας ARIMA(p, d, q) διαδικασίας θα είναι ίσο με:

$$\varphi(L)\Delta^d y_t = \theta(L)u_t$$

ή εναλλακτικά:

$$\varphi(L)(1 - L)^d y_t = \theta(L)u_t \quad (1.10)$$

όπου

$$\varphi(L) = (1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p) \text{ και}$$

$$\theta(L) = 1 - \theta_1 L^1 - \theta_2 L^2 - \dots - \theta_q L^q$$

Η εξίσωση (1.10) είναι γνωστή και ως εξίσωση διαφοράς (*difference equation form*) και η γενικότερη μεθοδολογία για την ανάπτυξη και κατασκευή υποδειγμάτων ARIMA είναι γνωστή ως μεθοδολογία των Box και Jenkins.

1.5 Δεδομένα Πάνελ

Τα δεδομένα πάνελ (*panel data - longitudinal data*) προέρχονται από το συνδυασμό διαστρωματικών δεδομένων και δεδομένων χρονοσειρών. Σε αντίθεση με τα διαστρωματικά δεδομένα και τα δεδομένα χρονοσειρών, τα δεδομένα πάνελ αφενός δίνουν τη δυνατότητα να μελετηθεί η κάθε μονάδα διαχρονικά και αφετέρου πολλές διαφορετικές μονάδες της ίδιας χρονικής περιόδου. Τα παραπάνω πλεονεκτήματα επιτρέπουν τη μελέτη δυναμικών πτυχών των προβλημάτων. Χρησιμοποιούνται ευρέως στις κοινωνικές και οικονομικές επιστήμες και διαρκώς κερδίζουν έδαφος έναντι των διαστρωματικών δεδομένων και των δεδομένων χρονοσειρών.

Η δημιουργία ενός συνόλου δεδομένων πάνελ απαιτεί τη λήψη παρατηρήσεων από τις ίδιες διαστρωματικές μονάδες (άτομα, οικογένειες, σχολικές μονάδες, επιχειρήσεις, δήμους, περιφέρειες κ.ά.) διαχρονικά. Για παράδειγμα ένα σύνολο δεδομένων πάνελ, που αφορούν τους μισθούς ατόμων, τις ώρες εργασίας, το επίπεδο εκπαίδευσης, καθώς και άλλους παράγοντες, συγκεντρώνεται με την τυχαία επιλογή προσώπων από έναν πληθυσμό σε μία δεδομένη χρονική στιγμή και στη συνέχεια από τα ίδια άτομα λαμβάνονται επανειλημμένες συνεντεύξεις σε διαδοχικές χρονικές στιγμές. Με αυτό τον τρόπο προκύπτουν δεδομένα για τους μισθούς, τις ώρες εργασίας, το επίπεδο

εκπαίδευσης και ούτω καθεξής, για την ίδια μονάδα ατόμων σε διαφορετικές χρονικές περιόδους. Επομένως, συμβολίζοντας τις παρατηρήσεις με x_{it} όπου $i = 1, 2, \dots, n$ η κάθε διαστρωματική μονάδα και $t = 1, \dots, T_i$ οι χρονικές περιόδους που έλαβε χώρα η έρευνα, τα δεδομένα θα έχουν την παρακάτω μορφή:

$$\begin{aligned} &1^{\text{η}} \text{ μονάδα ενδιαφέροντος} - \{x_{11}, x_{12}, \dots, x_{1T_1}\} \\ &2^{\text{η}} \text{ μονάδα ενδιαφέροντος} - \{x_{21}, x_{22}, \dots, x_{2T_2}\} \\ &\quad \vdots \\ &n^{\text{η}} \text{ μονάδα ενδιαφέροντος} - \{x_{n1}, x_{n2}, \dots, x_{nT_n}\} \end{aligned}$$

Ο δείκτης T_i , καθώς παρέχει την πληροφορία για το πλήθος των χρονικών περιόδων που έλαβε χώρα η έρευνα στην ουσία δείχνει το πλήθος των παρατηρήσεων που έχει συλλεχθεί για την κάθε διαστρωματική μονάδα. Στην περίπτωση που όλες οι μονάδες έλαβαν μέρος στην έρευνα σε όλες τις χρονικές περιόδους και άρα $T_1 = T_2 = \dots = T_n = T$ τότε το πάνελ δεδομένων ονομάζεται ισορροπημένο πάνελ (*balanced panel*) και το μέγεθος του δείγματος είναι ίσο με $N = nT$, ενώ αν το πλήθος των παρατηρήσεων διαφέρει από μονάδα σε μονάδα τότε το πάνελ δεδομένων ονομάζεται μη-ισορροπημένο (*unbalanced panel*) και το μέγεθος του δείγματος είναι ίσο με $T_1 + T_2 + \dots + T_n$.

Οφέλη και Μειονεκτήματα των Δεδομένων Πάνελ

Οι Hsiao (2003) και Klevmarken (1989) παραθέτουν τα οφέλη και τα μειονεκτήματα της χρήσης δεδομένων πάνελ όπως απαριθμούνται στη συνέχεια:

Οφέλη

- 1) Παρέχουν τη δυνατότητα ελέγχου των σταθερών χαρακτηριστικών, δηλαδή των χαρακτηριστικών που μεταβάλλονται μεταξύ των μονάδων ενδιαφέροντος, αλλά παραμένουν σταθερά διαχρονικά. Συνεπώς, μέσω αυτών μπορεί να ελεγχθεί η ετερογένεια (*heterogeneity*). Μονάδες όπως άτομα, εταιρίες, περιφέρειες, κράτη είναι ετερογενή. Η διαστρωματική ανάλυση και η ανάλυση χρονοσειρών δε μπορούν να ελέγξουν την ετερογένεια, με κίνδυνο τη λήψη μεροληπτικών αποτελεσμάτων.

- 2) Τα δεδομένα πάνελ δίνουν πιο κατατοπιστικά στοιχεία, μεγαλύτερη μεταβλητότητα, μικρότερη συγγραμικότητα μεταξύ των μεταβλητών, περισσότερους βαθμούς ελευθερίας και μεγαλύτερη αποτελεσματικότητα.
- 3) Με τα δεδομένα πάνελ δίνεται η δυνατότητα μελέτης δυναμικών φαινομένων, δηλαδή φαινομένων που μεταβάλλονται με το χρόνο. Αν για παράδειγμα ένας ερευνητής επιθυμεί να μελετήσει πχ. την εξαμηνιαία κερδοφορία ενός συνόλου επιχειρήσεων του κλάδου τροφίμων για μία περίοδο 5 ετών, τα δεδομένα πάνελ είναι τα πιο κατάλληλα, καθώς τα μεν διαστρωματικά δεδομένα δε μπορούν να εκφράσουν τέτοιου είδους δυναμικές σχέσεις, τα δε στοιχεία χρονοσειρών εκφράζουν δυναμικές σχέσεις, αλλά οι εκτιμήσεις που παράγουν δεν είναι ιδιαίτερα ακριβείς λόγω της ύπαρξης πολυσυγγραμικότητας.
- 4) 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 δεδομένα, όσο και κατά την εκτίμηση με στοιχεία χρονοσειρών. Αν για παράδειγμα, υπάρχει δείγμα 50 χωρών ($n = 50$) και ένας ερευνητής επιθυμεί να ελέγξει τις τιμές των κρατικών ομολόγων για μία περίοδο 10 ετών ($T = 10$), το πλήθος των παρατηρήσεων που θα χρησιμοποιήσει για την ανάλυση με δεδομένα πάνελ θα είναι $nT = 500$.
- 5) Δίνουν τη δυνατότητα υποδειματοποίησης μίας διαχρονική συμπεριφοράς που χαρακτηρίζει κάθε ξεχωριστή οντότητα i , αποφεύγοντας έτσι το σφάλμα προσθετικότητας διότι μία διαχρονική συμπεριφορά που χαρακτηρίζει κάποια ή κάποιες μονάδες ενδιαφέροντος, δε σημαίνει ότι θα πρέπει να χαρακτηρίζει και το σύνολο των μονάδων.

Μειονεκτήματα

Τα δεδομένα πάνελ παρέχουν πολλά προτερήματα σε σχέση με τα αμιγώς διαστρωματικά δεδομένα ή τα αμιγώς δεδομένα χρονοσειρών. Παρόλα αυτά συναντώνται αρκετές δυσκολίες στη διαδικασία της δημιουργίας τους και στην ανάλυσή τους, που παρουσιάζονται στη συνέχεια.

- 1) Ένα από τα πιο συνηθισμένα προβλήματα είναι αυτό της φθοράς (*attrition*) όπου κάποια ή κάποιες από τις μονάδες που μελετώνται αποχωρεί από το πείραμα πριν από τη λήξη του. Η φθορά μπορεί να οφείλεται στη χρεοκοπία κάποιας ή κάποιων επιχειρήσεων, στο θάνατο κάποιου ή κάποιων

ανθρώπων, στη συγχώνευση οργανισμών ή σχολικών μονάδων, στην αποστροφή κάποιων μονάδων του πειράματος για την έρευνα κλπ. Αν η φθορά έχει συστηματικό χαρακτήρα όπως για παράδειγμα αν από τον κλάδο τροφίμων αποχωρήσουν για κάποιο λόγο όλες οι εταιρίες που κατασκευάζουν ζυμαρικά, τότε υπάρχει εμφάνιση μεροληψίας επιλογής με αποτέλεσμα το σύνολο των δεδομένων πάνελ να χάνει ένα σημαντικό μέρος της αντιπροσωπευτικότητάς του και να μην μπορεί να χαρακτηριστεί τυχαίο.

- 2) Ύπαρξη σφαλμάτων μέτρησης, τα οποία μπορεί να οφείλονται σε ασαφείς ερωτήσεις, σε σκόπιμη διαστρέβλωση των απαντήσεων, σε κακή διαλογή των ερωτώμενων, στη μη καταγραφή των απαντήσεων από λάθος του ερευνητή, καθώς και σε διαφορετικές μονάδες μέτρησης.
- 3) Ύπαρξη σφαλμάτων επιλεκτικότητας (*selectivity problems*), τα οποία προκύπτουν είτε από τη μη απόκριση των επιλεγμένων διαστρωματικών μονάδων είτε από την υποκειμενικότητα της κάθε διαστρωματικής μονάδας

Παρόλο που η ύπαρξη δεδομένων πάνελ έδωσε τη δυνατότητα στην εμπειρική οικονομική έρευνα να ασχοληθεί με νέους τρόπους ανάλυσης της συμπεριφοράς των οικονομικών μονάδων, από την άλλη πλευρά έγειρε σημαντικά προβλήματα ως προς την χρήση των κατάλληλων οικονομετρικών υποδειγμάτων. Η ανάπτυξη οικονομετρικών υποδειγμάτων για την ανάλυση διαστρωματικών στοιχείων ξεκινάει από τις αρχές της δεκαετίας του 1960 αλλά παρόλα αυτά μόνο τα τελευταία 15 χρόνια παρουσιάζεται μία έντονη μεθοδολογική έρευνα που αποσκοπεί στην προσαρμογή και ανάπτυξη των οικονομετρικών υποδειγμάτων ικανών να αντιμετωπίσουν προβλήματα που αφορούν την εμπειρική εξειδίκευση, την οικονομετρική εκτίμηση και το στατιστικό έλεγχο οικονομικών υποδειγμάτων ανάλυσης της συμπεριφοράς των οικονομικών μονάδων. Ιδιαίτερα σημαντική προς την κατεύθυνση αυτή είναι η συμβολή των J.Heckman και D.McFadden οι οποίοι για τον λόγο αυτό τιμήθηκαν με το Νόμπελ Οικονομίας το 2000.

Ανάλυση Δεδομένων Πάνελ

Το ενδιαφέρον στην ανάλυση των δεδομένων πάνελ, εστιάζεται στους μη παρατηρούμενους παράγοντες που επηρεάζουν την εξαρτημένη μεταβλητή y_{it} και

είναι δύο ειδών: οι σταθεροί και οι διαχρονικά κυμαινόμενοι. Στην πιο απλή περίπτωση όπου το σύνολο δεδομένων πάνελ αποτελείται από δύο περιόδους, την αρχική περίοδο $t = 0$ και την τελική περίοδο $t = 1$ και μία μόνο ερμηνευτική μεταβλητή x_{it} το υπόδειγμα των απαρατήρητων ή σταθερών επιδράσεων είναι της μορφής:

$$y_{it} = \beta_0 + \delta_0 d_t + \beta_1 x_{it} + a_i + u_{it} \quad (1.11)$$

ή

$$y_{it} = \beta_0 + \delta_0 d_t + \beta_1 x_{it} + v_{it}, \quad (1.12)$$

όπου $v_{it} = a_i + u_{it}$.

Στο συμβολισμό y_{it} το i συμβολίζει το άτομο, την επιχείρηση, την πόλη και ούτω καθεξής ενώ το t τη χρονική περίοδο. Η μεταβλητή d_t είναι μία ψευδομεταβλητή που παίρνει την τιμή μηδέν όταν $t = 0$ και την τιμή ένα όταν $t = 1$. Δε μεταβάλλεται όταν μεταβάλλεται η διαστρωματική μονάδα και για αυτό το λόγο δεν έχει δείκτη i . Ο λόγος που συμπεριλαμβάνεται στο υπόδειγμα είναι ότι με αυτό τον τρόπο επιτρέπεται στην τεταγμένη της αρχής να μεταβάλλεται διαχρονικά, πράγμα πολύ χρήσιμο σε πληθώρα εφαρμογών. Επομένως, στην (1.12) η τεταγμένη της αρχής για $t = 0$ είναι β_0 ενώ για $t = 1$ είναι $\beta_0 + \delta_0$.

Η μεταβλητή a_i ονομάζεται είτε απαρατήρητη επίδραση (*unobserved effect*) είτε σταθερή επίδραση (*fixed effect*) είτε απαρατήρητη ετερογένεια (*unobserved heterogeneity*) και αποτυπώνει όλους τους μη παρατηρούμενους και διαχρονικά σταθερούς παράγοντες που επηρεάζουν την y_{it} . Ο λόγος που δεν έχει δείκτη t είναι γιατί παραμένει σταθερή διαχρονικά. Το σφάλμα u_{it} συχνά ονομάζεται σφάλμα ιδιοσυγκρασίας ή χρονικά μεταβαλλόμενο σφάλμα και αντιπροσωπεύει τους μη παρατηρούμενους παράγοντες που μεταβάλλονται διαχρονικά και επηρεάζουν την y_{it} . Τέλος, από το άθροισμα της απαρατήρητης επίδρασης a_i και του σφάλματος ιδιοσυγκρασίας u_{it} , προκύπτει το σύνθετο σφάλμα v_{it} .

Το βασικότερο ζήτημα, από το οποίο εξαρτάται η κατάλληλη επιλογή υποδείγματος και κατ' επέκταση η ορθή εκτίμησή του, είναι το αν η απαρατήρητη επίδραση a_i σχετίζεται με την ερμηνευτική μεταβλητή x_{it} . Στην περίπτωση που ισχύει κάτι τέτοιο, δηλαδή $Cov(a_i, x_{it}) \neq 0$, οι εκτιμήσεις των παραμέτρων μέσω των ελαχίστων τετραγώνων θα είναι μεροληπτικές και ασυνεπείς. Ως εκ τούτου η εξάλειψη των απαρατήρητων επιδράσεων, όταν αυτές σχετίζονται με την ερμηνευτική μεταβλητή, αποτελεί θέμα μέγιστης σημασίας. Στη συνέχεια περιγράφεται η εξίσωση πρώτων διαφορών και ο μετασχηματισμός των σταθερών επιδράσεων, μέθοδοι εξαιρετικά

χρήσιμες για την εξάλειψη των απαρατήρητων επιδράσεων, καθώς και το υπόδειγμα τυχαίων επιδράσεων που εφαρμόζεται όταν οι απαρατήρητες επιδράσεις δε σχετίζονται με τις ερμηνευτικές μεταβλητές του συνόλου των δεδομένων πάνελ.

Εξίσωση Πρώτων Διαφορών

Η εξάλειψη των απαρατήρητων επιδράσεων a_i καθώς και η εκτίμηση των παραμέτρων γίνεται αρχικά για τη σχέση (1.11) που αφορά μία απλή περίπτωση δεδομένων πάνελ δύο περιόδων με μία ερμηνευτική μεταβλητή. Για την κάθε περίοδο για την (1.11) θα ισχύει:

- για $t = 0$, $y_{i0} = \beta_0 + \beta_1 x_{i0} + a_i + u_{i0}$
- για $t = 1$, $y_{i1} = (\beta_0 + \delta_0) + \beta_1 x_{i1} + a_i + u_{i1}$

Αφαιρώντας από τη δεύτερη εξίσωση την πρώτη, προκύπτει:

$$y_{i1} - y_{i0} = \delta_0 + \beta_1(x_{i1} - x_{i0}) + (u_{i1} - u_{i0})$$

ή

$$\Delta y_i = \delta_0 + \beta_1 \Delta x_i + \Delta u_i \quad (1.13)$$

όπου το « Δ » συμβολίζει τη μεταβολή από $t = 0$ σε $t = 1$. Η σχέση (1.13) ονομάζεται εξίσωση πρώτων διαφορών (*first differencing equation*) και όπως γίνεται φανερό μέσω αυτής της μεθόδου η απαρατήρητη επίδραση εξαλείφεται, λόγω του ότι παραμένει σταθερή διαχρονικά. Ο σταθερός όρος δ_0 είναι στην ουσία η μεταβολή της τεταγμένης της αρχής από $t = 0$ σε $t = 1$.

Η εκτίμηση των παραμέτρων της εξίσωσης (1.13) μπορεί να γίνει με τη μέθοδο των ελαχίστων τετραγώνων αρκεί να πληρούνται οι βασικές προϋποθέσεις που αναφέρθηκαν στην παράγραφο (1.3). Η πιο σημαντική από αυτές είναι ότι η Δu_i δεν σχετίζεται με τη Δx_i . Η υπόθεση αυτή ισχύει αν το σφάλμα ιδιοσυγκρασίας, u_{it} , σε κάθε χρονική στιγμή t δε σχετίζεται με την ερμηνευτική μεταβλητή x_{it} . Ως εκ τούτου, οι ερμηνευτικές μεταβλητές πρέπει να είναι αυστηρά εξωγενείς. Μία ακόμη υπόθεση εξίσου σημαντική είναι ότι η Δx_i πρέπει να διαφοροποιείται κάπως διαχρονικά. Στην περίπτωση που $\Delta x_i = 0$ δεν μπορεί να επιτευχθεί εκτίμηση της (1.13) με τη μέθοδο ελαχίστων τετραγώνων. Τέλος, ο εκτιμητής της παραμέτρου β_1 που προκύπτει από τη μέθοδο ελαχίστων τετραγώνων είναι αμερόληπτος και συνεπής και ονομάζεται εκτιμητής πρώτων διαφορών.

Γενικεύοντας το σύνολο των δεδομένων πάνελ για T χρονικές περιόδους και k ερμηνευτικές μεταβλητές το υπόδειγμα των παρατηρήσιμων επιδράσεων θα είναι:

$$y_{it} = \delta_1 + \delta_2 d2_t + \dots + \delta_T dT_t + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + a_i + u_{it}$$

με την τεταγμένη της αρχής για την πρώτη περίοδο να είναι ίση με δ_1 , για τη δεύτερη περίοδο ίση με $\delta_1 + \delta_2$ και ούτω καθεξής. Για την εξάλειψη των παρατηρήσιμων επιδράσεων a_i αρκεί να αφαιρεθούν διαδοχικά η πρώτη περίοδος από τη δεύτερη, η δεύτερη από την τρίτη και ούτω καθεξής και με την υπόθεση ότι οι ερμηνευτικές μεταβλητές είναι αυστηρά εξωγενείς, δηλ. $Cov(x_{itj}, u_{is}) = 0, \forall t, s, j$, η εκτίμηση των αμερόληπτων και συνεπών παραμέτρων πραγματοποιείται μέσω της μεθόδου των ελαχίστων τετραγώνων.

Μετασχηματισμός των Σταθερών Επιδράσεων

Ο μετασχηματισμός των σταθερών επιδράσεων (*Fixed Effects Transformation*) είναι μία διαδικασία παρόμοιας φιλοσοφίας με αυτή των πρώτων διαφορών που μελετήθηκε στην προηγούμενη παράγραφο, κατά την οποία εξαλείφεται η παρατηρήσιμη επίδραση a_i , πριν από την εκτίμηση των παραμέτρων του υποδείγματος, καθώς και όλες οι χρονικά σταθερές ερμηνευτικές μεταβλητές.

Το υπόδειγμα των παρατηρήσιμων επιδράσεων για T χρονικές περιόδους και μία ερμηνευτική μεταβλητή x_{it} είναι της μορφής:

$$y_{it} = \beta_1 x_{it} + a_i + u_{it} \quad (1.14)$$

για $t = 1, 2, \dots, T$.

Λαμβάνοντας για κάθε διαστρωματική μονάδα i το μέσο όρο της εξίσωσης (1.14) διαχρονικά προκύπτει η σχέση:

$$\bar{y}_i = \beta_1 \bar{x}_i + a_i + \bar{u}_i \quad (1.15)$$

όπου $\bar{y}_i = \frac{1}{T} \sum_{t=1}^T y_{it}$ και ούτω καθεξής.

Αφαιρώντας κατά μέλη τις σχέσεις (1.14) και (1.15) προκύπτει:

$$y_{it} - \bar{y}_i = \beta_1 (x_{it} - \bar{x}_i) + u_{it} - \bar{u}_i$$

ή

$$\dot{y}_{it} = \beta_1 \dot{x}_{it} + \dot{u}_{it} \quad (1.16)$$

για $t = 1, 2, \dots, T$, όπου η $\dot{y}_{it} = y_{it} - \bar{y}_i$ είναι τα χρονικά προσαρμοσμένα δεδομένα (*time demeaned data*) για την y και ομοίως για την \dot{x}_{it} και την \dot{u}_{it} . Ο μετασχηματισμός σταθερών επιδράσεων ονομάζεται και μετασχηματισμός εντός (*within transformation*).

Με αυτή τη διαδικασία εξαλείφεται η απαρατήρητη επίδραση a_i και είμαστε σε θέση να εκτιμήσουμε η σχέση (1.16) μέσω των ομαδοποιημένων ελαχίστων τετραγώνων (*pooled ordinary least squares*) ομαδοποιώντας απλώς τις διαστρωματικές μονάδες ανά χρονική περίοδο. Ο εκτιμητής των ομαδοποιημένων ελαχίστων τετραγώνων που στηρίζεται στις χρονικά προσαρμοσμένες μεταβλητές ονομάζεται εκτιμητής σταθερών επιδράσεων (*fixed effects estimator*) ή εκτιμητής εντός (*within estimator*).

Γενικεύοντας την (1.14) με την προσθήκη k ερμηνευτικών μεταβλητών, το υπόδειγμα των απαρατήρητων επιδράσεων είναι της μορφής:

$$y_{it} = \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + a_i + u_{it}$$

για $t = 1, 2, \dots, T$.

Χρησιμοποιώντας τη χρονική προσαρμογή σε κάθε ερμηνευτική μεταβλητή και πραγματοποιώντας μία παλινδρόμηση ομαδοποιημένων ελαχίστων τετραγώνων προκύπτει η γενική χρονικά προσαρμοσμένη εξίσωση η οποία είναι:

$$\dot{y}_{it} = \beta_1 \dot{x}_{it1} + \beta_2 \dot{x}_{it2} + \dots + \beta_k \dot{x}_{itk} + \dot{u}_{it}$$

για $t = 1, 2, \dots, T$.

Όπως στην εξίσωση των πρώτων διαφορών, έτσι και εδώ, μία πολύ βασική υπόθεση για τη λήψη αμερόληπτων εκτιμητών είναι αυτή των αυστηρά εξωγενών μεταβλητών όπου το σφάλμα ιδιοσυγκρασίας δεν πρέπει να σχετίζεται με καμία ερμηνευτική μεταβλητή σε κάθε χρονική περίοδο, δηλ. $Cov(x_{itj}, u_{is}) = 0, \forall t, s, j$.

Υπόδειγμα Τυχαίων Επιδράσεων

Το υπόδειγμα τυχαίων επιδράσεων (*Random effects model*) χρησιμοποιείται όταν υποθέσουμε ότι η απαρατήρητη επίδραση a_i , δε σχετίζεται με κάθε ερμηνευτική μεταβλητή. Συνεπώς:

$$Cov(x_{itj}, a_i) = 0 \quad (1.17)$$

Το υπόδειγμα των απαρατήρητων επιδράσεων είναι της μορφής:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + a_i + u_{it} \quad (1.18)$$

στο οποίο συμπεριλαμβάνεται η τεταγμένη της αρχής β_0 ώστε να είναι εφικτή η διατύπωση της υπόθεσης ότι η απαρατήρητη επίδραση a_i έχει μηδενικό μέσο. Συνήθως μεταξύ των ερμηνευτικών μεταβλητών επιτρέπεται να υπάρχουν και χρονικές ψευδομεταβλητές. Όταν χρησιμοποιείται η μέθοδος των σταθερών επιδράσεων ή η μέθοδος των πρώτων διαφορών, ο στόχος μας είναι να εξάλειψη της a_i καθώς

θεωρείται ότι σχετίζεται με μία ή περισσότερες από τις x_{itj} . Στο υπόδειγμα τυχαίων επιδράσεων όμως η εξάλειψη αυτή οδηγεί σε μη αποτελεσματικούς εκτιμητές.

Σύμφωνα με τη σχέση (1.17) μπορεί να θεωρηθεί ότι η εκτίμηση του υποδείγματος (1.18) μπορεί να επιτευχθεί είτε μέσω της ανάλυσης παλινδρόμησης για απλά διαστρωματικά δεδομένα είτε μιας ομαδοποιημένης διαδικασίας ελαχίστων τετραγώνων, χωρίς τη χρήση δεδομένων πάνελ. Στην πρώτη περίπτωση δε λαμβάνονται υπόψη χρήσιμες πληροφορίες για τις χρονικές περιόδους, ενώ στη δεύτερη περίπτωση δε λαμβάνεται υπόψη ένα πολύ βασικό χαρακτηριστικό του υποδείγματος. Ορίζοντας το σύνθετο σφάλμα $v_{it} = a_i + u_{it}$, η σχέση (1.18) γίνεται:

$$y_{it} = \beta_0 + \beta_1 x_{it1} + \beta_2 x_{it2} + \dots + \beta_k x_{itk} + v_{it}$$

και οι v_{it} παρουσιάζουν διαχρονικά αυτοσυσχέτιση, λόγω του ότι η a_i βρίσκεται στο σύνθετο σφάλμα σε κάθε χρονική περίοδο. Πιο συγκεκριμένα αποδεικνύεται ότι

$$\text{Corr}(v_{it}, v_{is}) = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_u^2}, t \neq s$$

όπου $\sigma_a^2 = \text{Var}(a_i)$ και $\sigma_u^2 = \text{Var}(u_{it})$. Αν δεν ληφθεί σοβαρά στα υπόψη η αυτοσυσχέτιση και το υπόδειγμα εκτιμηθεί με τη μέθοδο ελαχίστων τετραγώνων οι εκτιμητές που θα προκύψουν θα είναι ασυνεπείς. Για το λόγο αυτό η πιο κατάλληλη μέθοδος είναι αυτή των γενικευμένων ελαχίστων τετραγώνων (*Generalized Least Squares – GLS*) της οποίας το μαθηματικό μέρος είναι αρκετά πολύπλοκο. Στην πράξη επιδιώκεται η απαλοιφή των αυτοσυσχετίσεων για την οποία απαιτείται μεγάλο N και

σχετικά μικρό T . Ορίζοντας $\lambda = 1 - \left[\frac{\sigma_u^2}{\sigma_u^2 + T\sigma_a^2} \right]^{\frac{1}{2}} \in (0,1)$ προκύπτει η μετασχηματισθείσα εξίσωση των τυχαίων επιδράσεων:

$$y_{it} - \lambda \bar{y}_i = \beta_0(1 - \lambda) + \beta_1(x_{it1} - \lambda \bar{x}_{i1}) + \dots + \beta_k(x_{itk} - \lambda \bar{x}_{ik}) + \beta_1(v_{it} - \lambda \bar{v}_i),$$

όπου η παύλα πάνω από τα γράμματα συμβολίζει όπως και στο υπόδειγμα σταθερών επιδράσεων, τους χρονικούς μέσους. Η εκτίμηση $\hat{\lambda}$ της παραμέτρου λ ονομάζεται εκτιμητής τυχαίων επιδράσεων (*random effects estimator*), είναι συνεπής αλλά όχι αμερόληπτος και ασυμπτωτικά κατανέμεται κανονικά καθώς το N αυξάνεται και το T παραμένει σταθερό.

Από εμπειρικές εργασίες διαπιστώνεται ότι οι συγγραφείς επιλέγουν ανάμεσα στις σταθερές και τις τυχαίες επιδράσεις με βάση το κατά πόσο οι απαραίτητες επιδράσεις a_i θεωρούνται παράμετροι προς εκτίμηση ή αποτελέσματα μίας τυχαίας μεταβλητής. Όταν δε μπορούν να θεωρηθούν οι παρατηρήσεις τυχαίες λήψεις από ένα μεγάλο πληθυσμό, οι απαραίτητες επιδράσεις αντιμετωπίζονται ως παράμετροι που πρέπει

να εκτιμηθούν και έτσι χρησιμοποιείται η μέθοδος των σταθερών επιδράσεων. Ακόμα και αν ο ερευνητής αποφασίσει να αντιμετωπίσει τις απαρατήρητες επιδράσεις a_i ως τυχαίες μεταβλητές, πρέπει να ελέγξει το κατά πόσο είναι ασυσχέτιστες με τις ερμηνευτικές μεταβλητές. Αν επικρατεί η υπόθεση ότι οι a_i δε σχετίζονται με όλες τις ερμηνευτικές μεταβλητές, τότε η μέθοδος των τυχαίων επιδράσεων είναι απαραίτητη. Από την άλλη μεριά, αν οι a_i σχετίζονται με κάποιες ερμηνευτικές μεταβλητές, η μέθοδος των σταθερών επιδράσεων είναι απαραίτητη. Αν χρησιμοποιηθεί η μέθοδος των τυχαίων επιδράσεων, τότε οι εκτιμητές είναι κατά κανόνα ασυνεπείς. Επίσης πρέπει να θεωρηθεί δεδομένο πως σε σύνολα δεδομένων πάνελ που ο ερευνητής επιθυμεί να αξιολογήσει τις επιπτώσεις μίας πολιτικής ακολουθείται πάντα το υπόδειγμα σταθερών επιδράσεων.

Η σύγκριση των εκτιμήσεων των τυχαίων επιδράσεων και των εκτιμήσεων των σταθερών επιδράσεων, μπορεί να αποτελέσει έναν έλεγχο για να διαπιστωθεί αν υπάρχει συσχέτιση ανάμεσα στις απαρατήρητες επιδράσεις και στις ερμηνευτικές μεταβλητές, υποθέτοντας ότι τα σφάλματα ιδιοσυγκρασίας και οι ερμηνευτικές μεταβλητές είναι ασυσχέτιστες σε όλες τις χρονικές περιόδους. Ο πρώτος που πρότεινε αυτόν τον έλεγχο είναι ο Hausman (1978).

1.6 Ανακεφαλαίωση

Στα πλαίσια του παρόντος κεφαλαίου έγινε προσπάθεια να τεθεί με επαγωγικό τρόπο το θεωρητικό υπόβαθρο για την ανάλυση των επόμενων κεφαλαίων και πιο συγκεκριμένα της μεθόδου των διαφορών σε διαφορές. Ειδικότερα, έγινε αναφορά στις σημαντικότερες κατηγορίες συνόλων δεδομένων που αποτελούν αντικείμενο μελέτης της οικονομετρίας καθώς και τα σημαντικότερα μαθηματικά υποδείγματα που χρησιμοποιούνται για τη υποδειματοποίηση αυτών, όπως η ανάλυση παλινδρόμησης, η ανάλυση χρονοσειρών και η ανάλυση δεδομένων πάνελ. Αναφέρθηκαν προβλήματα που ανακύπτουν στην έλλειψη στασιμότητας μίας χρονοσειράς και μέθοδοι αντιμετώπισής τους, καθώς και οι τρόποι αντιμετώπισης των απαρατήρητων επιδράσεων σε υποδείγματα δεδομένων πάνελ.

ΚΕΦΑΛΑΙΟ 2

ΜΕΘΟΔΟΣ ΔΙΑΦΟΡΩΝ ΣΕ ΔΙΑΦΟΡΕΣ

2.1 Εισαγωγή

Στην εμπειρική ανάλυση, απότομες μεταβολές (γνωστές και ως *θεραπεία - Treatment*) στο οικονομικό περιβάλλον, στην κυβερνητική πολιτική και στο θεσμικό περιβάλλον, συνθέτουν μία πολύ ενδιαφέρουσα κατηγορία πειραμάτων, τα φυσικά πειράματα ή εναλλακτικά ονομάζονται πειράματα (*quasi experiments*). Το φυσικό πείραμα λαμβάνει χώρα όταν κάποιο εξωγενές περιστατικό μεταβάλλει το περιβάλλον στο οποίο δραστηριοποιούνται άτομα, οικογένειες, επιχειρήσεις, πόλεις, κράτη και αποτελείται πάντα από μία ομάδα ελέγχου (*Control Group*), που δεν επηρεάζεται από τη μεταβολή της πολιτικής και μία ομάδα αντιμετώπισης ή θεραπείας (*Treatment Group*) που πιστεύεται ότι επηρεάζεται από τη μεταβολή αυτή. Σε αντίθεση με ένα ελεγχόμενο πείραμα, όπου η ομάδα αντιμετώπισης και η ομάδα ελέγχου επιλέγονται τυχαία και με σαφή τρόπο, στα φυσικά πειράματα οι ομάδες αυτές προκύπτουν από τη συγκεκριμένη μεταβολή της πολιτικής. Για να ληφθούν υπόψη οι συστηματικές διαφορές ανάμεσα στην ομάδα ελέγχου και στην ομάδα αντιμετώπισης απαιτούνται δεδομένα δύο περιόδων, μίας περιόδου πριν από τη μεταβολή και μίας περιόδου μετά από αυτήν. Ως εκ τούτου το δείγμα αναλύεται σε τέσσερις ομάδες: στην ομάδα ελέγχου πριν από τη μεταβολή, στην ομάδα ελέγχου μετά τη μεταβολή, στην ομάδα αντιμετώπισης πριν τη μεταβολή και στην ομάδα αντιμετώπισης μετά τη μεταβολή. Η μέθοδος των διαφορών σε διαφορές (*Difference-in-Differences*) χρησιμοποιείται για να αξιολογήσει και να ερμηνεύσει τις επιδράσεις των μεταβολών αυτών, στην ομάδα αντιμετώπισης.

Στο παρόν κεφάλαιο παρουσιάζεται και αναλύεται η μέθοδος των διαφορών σε διαφορές. Στην αρχή γίνεται αναφορά στα προβλήματα της εσωτερικής και εξωτερικής εγκυρότητας που είναι δυνατό να προκύψουν κατά την ανάλυση των δεδομένων καθώς και στους εκτιμητές απλών διαφορών (*single difference estimators*) με στόχο να αναδειχθούν οι αδυναμίες τους. Στη συνέχεια η προσοχή εστιάζεται στους οι εκτιμητές των διαφορών σε διαφορές (*difference in differences estimators*), στο πως αυτοί ξεπερνούν τις αδυναμίες αυτές, καθώς και στο πώς ελέγχεται η εγκυρότητα τους.

2.2 Προβλήματα στην Εγκυρότητα

Είναι πολύ σημαντικό στο σημείο αυτό να αναφερθούν κάποια γενικά προβλήματα που προκύπτουν στην εξαγωγή των συμπερασμάτων στις εμπειρικές μελέτες και μπορούν να δώσουν τελείως διαφορετική ερμηνεία από αυτή που πραγματικά ισχύει. Τα προβλήματα στην εγκυρότητα μπορεί να είναι δύο μορφών. Προβλήματα εσωτερικής εγκυρότητας και προβλήματα εξωτερικής εγκυρότητας. Η εσωτερική εγκυρότητα αναφέρεται στο αν οι μεταβολές στην εξαρτημένη μεταβλητή οφείλονται στις μεταβολές των ερμηνευτικών μεταβλητών και μόνο σε αυτές, ενώ η εξωτερική εγκυρότητα αναφέρεται στο αν τα αποτελέσματα που προκύπτουν από μία μελέτη μπορούν να γενικευθούν σε διαφορετικά άτομα και περιβάλλοντα.

Προβλήματα Εσωτερικής Εγκυρότητας

Η εσωτερική εγκυρότητα μπορεί να παραβιαστεί από:

- *Παραλειπόμενες μεταβλητές*: Γεγονότα που επηρεάζουν τις παρατηρήσεις πριν και μετά από την παρέμβαση και δεν έχουν συμπεριληφθεί στη μελέτη.
- *Χρονολογικές τάσεις*: Μεταβολές στις μονάδες συναρτήσεως του χρόνου, όπως ο πληθωρισμός και η αύξηση των μισθών.
- *Κακώς προσδιορισμένες διακυμάνσεις* που προκαλούν υπερεκτίμηση των στατιστικών ελέγχων.
- *Σφάλμα μέτρησης* που προκύπτει από αλλαγές στην ορολογία ή στη μέθοδο της έρευνας και μπορεί να προκαλέσει αλλαγές στη μέτρηση των μεταβλητών. Αυτό μπορεί να περιλαμβάνει αλλαγές στη διατύπωση της έρευνας ή αλλαγή στη σειρά των ερωτήσεων.
- *Ενδογένεια* των ερμηνευτικών μεταβλητών.
- *Σφάλμα επιλεκτικότητας*: Αναφέρεται στις περιπτώσεις εκείνες όπου το χρησιμοποιούμενο δείγμα δεν επιλέγεται τυχαία από τον πληθυσμό.
- *Φθορά*: Είναι η απώλεια μονάδων από το πείραμα λόγω χρεοκοπίας, θανάτου και ούτω καθεξής.
- *Παραλειπόμενες αλληλεπιδράσεις*: Φανερόνουν διαφορετικές τάσεις στις ομάδες ελέγχου και αντιμετώπισης.

Προβλήματα Εξωτερικής Εγκυρότητας

Τα προβλήματα στην εξωτερική εγκυρότητα, έχουν να κάνουν με την πιθανότητα να υπάρχουν σημαντικές αλληλεπιδράσεις μεταξύ της εφαρμογής της μεταβολής της πολιτικής και των ατομικών χαρακτηριστικών, της τοποθεσίας και του χρόνου, τέτοιες ώστε τα αποτελέσματα μίας μελέτης να μη μπορούν να γενικευθούν σε διαφορετικές μονάδες και περιβάλλοντα. Πιο συγκεκριμένα οι αλληλεπιδράσεις αυτές είναι:

- *Αλληλεπίδραση της εφαρμογής της πολιτικής και της ομάδας αντιμετώπισης:* Η ομάδα αντιμετώπισης δεν είναι αντιπροσωπευτική του πληθυσμού.
- *Αλληλεπίδραση της εφαρμογής της πολιτικής και της τοποθεσίας:* Η επίδραση της παρέμβασης μπορεί να διαφέρει από περιοχή σε περιοχή ή από θεσμικό πλαίσιο σε θεσμικό πλαίσιο.
- *Αλληλεπίδραση της εφαρμογής της πολιτικής και του χρόνου:* Η επίδραση της παρέμβασης μπορεί να διαφέρει σε διαφορετικές χρονικές περιόδους.

Το προβλήματα της εγκυρότητας που αναφέρθηκαν, εσωτερικής ή εξωτερικής, είναι πιθανό να προκύψουν σε μία εμπειρική μελέτη, αλλά όχι δεδομένο. Η αντιμετώπιση αυτών στηρίζεται στην πείρα και στο θεωρητικό υπόβαθρο του εκάστοτε ερευνητή και όχι σε κάποιες συγκεκριμένες διαδικασίες που πρέπει ακολουθηθούν.

2.3 Εκτιμητές Απλών Διαφορών

Η αξιολόγηση των επιπτώσεων ενός φυσικού πειράματος μπορεί να πραγματοποιηθεί πέρα από τη μέθοδο των διαφορών σε διαφορές και με δύο άλλες πιο απλές μεθόδους, όπου η πρώτη συγκρίνει τις τιμές της ομάδας αντιμετώπισης με αυτές της ομάδας ελέγχου για την περίοδο μετά τη μεταβολή της πολιτικής και η δεύτερη της τιμές της ομάδας αντιμετώπισης πριν και μετά τη μεταβολή. Οι μέθοδοι αυτοί παρουσιάζουν αρκετές αδυναμίες και για αυτό το λόγο χρησιμοποιούνται για να δώσουν μία πιο διαισθητική εικόνα όταν δεν υπάρχουν δεδομένα και από τις τέσσερις ομάδες. Οι μέθοδοι αυτοί καθώς και οι όποιες αδυναμίες τους παρουσιάζονται στη συνέχεια.

Σύγκριση Ομάδων Μίας Περιόδου

Η εκτίμηση της επίδρασης από την εφαρμογή μίας πολιτικής (*Treatment Effect*), με δεδομένα μίας ή περισσότερων περιόδων μετά την εφαρμογή αυτής, πραγματοποιείται με τη σύγκριση των αποτελεσμάτων της ομάδας αντιμετώπισης και της ομάδας ελέγχου. Η μέθοδος αυτή συνήθως χρησιμοποιείται, όταν δεν υπάρχουν διαθέσιμα δεδομένα πριν από τη μεταβολή. Εφαρμόζοντας τα παραπάνω στην πράξη, για ένα σύνολο διαστρωματικών δεδομένων με μεταβλητές, τη μεταβλητή απόκρισης y και την ψευδομεταβλητή d , η οποία λαμβάνει την τιμή μηδέν αν η διαστρωματική μονάδα i ανήκει στην ομάδα ελέγχου και την τιμή ένα αν ανήκει στην ομάδα αντιμετώπισης, το υπόδειγμα που χρησιμοποιείται για τη σύγκριση των δύο ομάδων προκύπτει από τη μέθοδο ελαχίστων τετραγώνων και είναι της μορφής:

$$y = \beta_0 + \beta_1 d + u \quad (2.1)$$

ή

$$E(y|d) = \beta_0 + \beta_1 d$$

Δεδομένου ότι αυτή είναι μία απλή παλινδρόμηση επί μίας μόνο ψευδομεταβλητής, η τεταγμένη της αρχής είναι η μέση τιμή των τιμών της ομάδας ελέγχου, ενώ ο συντελεστής β_1 της ψευδομεταβλητής, είναι η διαφορά στις μέσες τιμές των δύο ομάδων. Επομένως, η αναμενόμενη τιμή της μεταβλητής απόκρισης για τις παρατηρήσεις που ανήκουν σε κάθε μία από τις δύο ομάδες θα είναι:

- Ομάδα ελέγχου:

$$E(y|d = 0) = \beta_0 \quad (2.2)$$

- Ομάδα αντιμετώπισης:

$$E(y|d = 1) = \beta_0 + \beta_1 \quad (2.3)$$

Αφαιρώντας από τη σχέση (2.3) τη σχέση (2.2) προκύπτει η επίδραση της επιβολής της νέας πολιτικής, στην ομάδα θεραπείας η οποία είναι ίση με το συντελεστή β_1 όπως φαίνεται και από την επόμενη σχέση:

$$E(y|d = 1) - E(y|d = 0) = \beta_1 \quad (2.4)$$

Υπάρχουν περιπτώσεις όπου τα δεδομένα αφορούν περισσότερες από μία περιόδους μετά την μεταβολή, σχηματίζοντας με αυτόν τον τρόπο ένα σύνολο δεδομένων πάνελ. Για την εκτίμηση της επίδρασης της μεταβολής αυτής αγνοείται συνειδητά η ύπαρξη χρονοσειράς για κάθε διαστρωματική μονάδα δημιουργώντας εκ νέου ένα σύνολο δεδομένων μονής διαστρωμάτωσης, υπολογίζοντας το μέσο όρο των παρατηρήσεων

της κάθε διαστρωματικής μονάδας για όλες τις χρονικές περιόδους. Στη συνέχεια μέσω της μεθόδου ελαχίστων τετραγώνων λαμβάνεται η εξίσωση του υποδείγματος η οποία είναι όμοια με τη σχέση (2.1) με τη μόνη διαφορά ότι η μεταβλητή απόκρισης αναφέρεται πλέον σε μέσες τιμές. Η μέθοδος είναι αξιόπιστη μόνο αν η μεταβολή έχει επέλθει ταυτόχρονα σε όλες τις διαστρωματικές μονάδες που ανήκουν στην ομάδα αντιμετώπισης. Εναλλακτικά η ύπαρξη πολλών περιόδων μπορεί να αντιμετωπιστεί με τη χρήση ψευδομεταβλητών περιόδων. Πιο συγκεκριμένα μπορεί κάποιος να τροποποιήσει τη σχέση (2.1) επιτρέποντας στην επίδραση του φυσικού πειράματος να μεταβάλλεται με το χρόνο, αλληλεπιδρώντας την ψευδομεταβλητή της κάθε περιόδου με την ψευδομεταβλητή d . Ως εκ τούτου για T χρονικές περιόδους η εξίσωση (2.1) γίνεται:

$$y = \beta_0 + \beta_1 d \times p_1 + \beta_2 d \times p_2 + \dots + \beta_T d \times p_T + u$$

όπου p_t , $t = 1, 2, \dots, T$ η ψευδομεταβλητή της κάθε χρονικής περιόδου, η οποία λαμβάνει την τιμή ένα αν η παρατήρηση ανήκει στη χρονική περίοδο t και μηδέν σε διαφορετική περίπτωση. Στον Πίνακα 2.1 παρουσιάζονται οι επιδράσεις του φυσικού πειράματος στην ομάδα αντιμετώπισης σε κάθε χρονική περίοδο.

Πίνακας 2.1

Επιδράσεις φυσικού πειράματος σε πολλαπλές περιόδους

	$t=1$	$t=2$	$t=3$...	$t=T$
<i>Ομάδα Αντιμετώπισης</i>	$\beta_0 + \beta_1$	$\beta_0 + \beta_2$	$\beta_0 + \beta_3$...	$\beta_0 + \beta_T$
<i>Ομάδα Ελέγχου</i>	β_0	β_0	β_0	...	β_0
<i>Επίδραση Μεταβολής</i>	β_1	β_2	β_3	...	β_T

Το πρόβλημα που παρουσιάζεται στη μέθοδο αυτή, είναι ότι ο ερευνητής δεν είναι σε θέση να γνωρίζει αν υπάρχει συστηματική διαφορά στην ομάδα ελέγχου και στην ομάδα αντιμετώπισης, η οποία δεν έχει καταγραφεί λόγω έλλειψης δεδομένων. Επομένως, η εκτιμηθείσα επίδραση είναι πιθανό να μην αντικατοπτρίζει την πραγματικότητα και να υπάρχει υπερεκτίμηση της επίδρασης της πολιτικής. Για παράδειγμα έστω ότι κάποιος ενδιαφέρεται να μελετήσει την επίδραση ενός προγράμματος επαγγελματικής κατάρτισης και βελτίωσης δεξιοτήτων, στις μηνιαίες αποδοχές των εργαζομένων. Η ομάδα ελέγχου είναι τα άτομα που δεν έλαβαν μέρος στο πρόγραμμα αυτό, ενώ η ομάδα αντιμετώπισης οι εργαζόμενοι που έλαβαν μέρος.

Συνηθίζεται τα άτομα που λαμβάνουν μέρος σε προγράμματα επαγγελματικής κατάρτισης να είναι ούτως ή άλλως πιο πρόθυμα να εργαστούν περισσότερο και σκληρότερα, οπότε έχουν μεγαλύτερες απολαβές από τους μη συμμετέχοντες ακόμα και χωρίς προγράμματα κατάρτισης.

Σύγκριση Μίας Ομάδας σε Δύο Περιόδους

Ένας δεύτερος τρόπος εκτίμησης της επίδρασης μίας νέας πολιτικής, είναι η σύγκριση των αποτελεσμάτων της περιόδου μετά την εφαρμογή της (*post-treatment*) με αυτά πριν από την εφαρμογή της (*pre-treatment*) μόνο για την ομάδα αντιμετώπισης. Αυτή η προσέγγιση χρησιμοποιείται ευρέως στον κλάδο των χρηματοοικονομικών αλλά δεν παύει να έχει κάποιες αδυναμίες.

Οι Bertrand, Schoar, και Thesmar (2007) εξετάζουν τον αντίκτυπο της απελευθέρωσης του τραπεζικού συστήματος της Γαλλίας, στη συμπεριφορά των δανειοληπτών και των τραπεζών. Οι Blanchard, Lopez-de-Silanes, και Shleifer (1994) συγκρίνουν ένα πλήθος εταιρικών δεικτών (όπως επενδύσεις, μερίσματα, πωλήσεις περιουσιακών στοιχείων κ.α.) πριν και μετά από μεγάλες νομικές επιβραβεύσεις. Οι Khwaja και Mian (2008) χρησιμοποιούν τις μη αναμενόμενες πυρηνικές δοκιμές στο Πακιστάν, για να εξετάσουν τις επιπτώσεις της έλλειψης οικονομικής ρευστότητας στην πολιτική δανεισμού των τραπεζών. Σε κάθε μία από τις παραπάνω περιπτώσεις οι ομάδες ελέγχου και αντιμετώπισης καθορίζονται από τις περιόδους πριν και μετά από το κάθε συμβάν, αποτελούμενες από τις ίδιες διαστρωματικές μονάδες.

Το σύνολο δεδομένων που προκύπτει, περιλαμβάνει δύο παρατηρήσεις για κάθε διαστρωματική μονάδα, μία πριν από το συμβάν και μία μετά από αυτό, το οποίο δεν είναι τίποτα άλλο παρά ένα σύνολο δεδομένων πάνελ δύο περιόδων. Η σύγκριση των αποτελεσμάτων επιτυγχάνεται με μία παλινδρόμηση δεδομένων πάνελ δύο περιόδων με την εξίσωση παλινδρόμησης να είναι της μορφής:

$$y = \beta_0 + \beta_1 p + u \quad (2.5)$$

ή

$$E(y|p) = \beta_0 + \beta_1 p$$

όπου η μεταβλητή y αντιστοιχεί στη μεταβλητή απόκρισης και η μεταβλητή p είναι μία ψευδομεταβλητή που λαμβάνει την τιμή μηδέν για τις παρατηρήσεις πριν από το συμβάν και την τιμή ένα για τις παρατηρήσεις μετά το συμβάν. Δεδομένου ότι αυτή

είναι μία απλή παλινδρόμηση επί μίας μόνο ψευδομεταβλητής, η τεταγμένη της αρχής είναι η μέση τιμή των τιμών της περιόδου πριν την πραγματοποίηση της μεταβολής, ενώ ο συντελεστής β_1 της ψευδομεταβλητής p , είναι η διαφορά στις μέσες τιμές των δύο περιόδων. Συνοψίζοντας τα παραπάνω προκύπτει:

- Πριν από το συμβάν:

$$E(y|p = 0) = \beta_0 \quad (2.6)$$

- Μετά από το συμβάν:

$$E(y|p = 1) = \beta_0 + \beta_1 \quad (2.7)$$

Αφαιρώντας από τη σχέση (2.7) τη σχέση (2.6) προκύπτει η επίδραση της νέας πολιτικής, στην ομάδα θεραπείας η οποία είναι ίση με το συντελεστή β_1 όπως φαίνεται παρακάτω:

$$E(y|p = 1) - E(y|p = 0) = \beta_1$$

Για περισσότερες χρονικές περιόδους πριν και μετά από το συμβάν υπάρχουν αρκετές διαθέσιμες επιλογές, ανάλογα με το αν θεωρηθεί ότι οι επιδράσεις της νέας πολιτικής είναι σταθερές για όλες τις περιόδους, ή αν μεταβάλλονται. Στην πρώτη περίπτωση ο ερευνητής οδηγείται σε ομαδοποίηση των δεδομένων πάνελ πολλών περιόδων σε δεδομένα πάνελ δύο περιόδων. Το υπόδειγμα που ερμηνεύει τις ζητούμενες επιδράσεις είναι το ίδιο με το (2.5) δηλαδή:

$$y = \beta_0 + \beta_1 p + u$$

Εναλλακτικά στη δεύτερη περίπτωση που οι επιδράσεις διαφέρουν από περίοδο σε περίοδο, η ύπαρξη δεδομένων πολλών περιόδων αντιμετωπίζεται με τη δημιουργία $T - 1$ ψευδομεταβλητών περιόδων για το σύνολο των T περιόδων έτσι ώστε να μη παρουσιαστεί το πρόβλημα της πολυσυγγραμικότητας. Η εξίσωση του υποδείγματος θα έχει τη μορφή:

$$y = \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \dots + \beta_{T-1} p_{T-1} + u,$$

όπου $p_s, s = 1, 2, \dots, T - 1$ ισούται με 1 για την περίοδο s και μηδέν για οποιαδήποτε άλλη περίοδο. Με τους εκτιμώμενους συντελεστές μπορεί κανείς να ελέγξει είτε οπτικά, είτε στατιστικά αν υπάρχει κάποια σημαντική αλλαγή γύρω από την ημερομηνία του φυσικού πειράματος. Ωστόσο, αν η εφαρμογή της νέας πολιτικής δεν πραγματοποιηθεί ταυτόχρονα σε όλες τις διαστρωματικές μονάδες, η στρατηγική αυτή δεν έχει τα επιθυμητά αποτελέσματα.

Προβλήματα που ανακύπτουν από τη μέθοδο αυτή σχετίζονται με τις χρονικές τάσεις που ενδεχομένως παρουσιάζουν τα δεδομένα, όπως η κυκλικότητα. Για

παράδειγμα ένας ερευνητής ενδιαφέρεται να μελετήσει την επίδραση της επαγγελματικής κατάρτισης στην αποδοτικότητα των εργαζομένων για μία εταιρεία που παράγει παγωτά. Αν η κατάρτιση πραγματοποιηθεί στα τέλη της καλοκαιρινής περιόδου, είναι δεδομένο ότι μετά από αυτήν οι πωλήσεις θα αρχίσουν σταδιακά να μειώνονται λόγω των γενικότερων καταναλωτικών προτιμήσεων. Κάτι τέτοιο θα οδηγήσει σε υποεκτίμηση της επίδρασης της κατάρτισης.

Για την αντιμετώπιση των αδυναμιών που εμφανίζονται μπορεί κάποιος να εξετάσει παράλληλα εταιρίες που δραστηριοποιούνται στον ίδιο κλάδο αλλά οι εργαζόμενοι τους δεν έλαβαν μέρος σε κανένα πρόγραμμα κατάρτισης. Με αυτόν τον τρόπο δημιουργείται καλύτερη εικόνα της επίδρασης του προγράμματος μιας και αναμένεται να υπάρχουν απότομες μεταβολές στην εταιρία που έλαβε μέρος στο πρόγραμμα σε σχέση με αυτές που δεν έλαβαν μέρος. Στην πραγματικότητα αυτή η προσέγγιση οδηγεί στη μέθοδο των διαφορών σε διαφορές η οποία και ακολουθεί στη συνέχεια.

2.4 Εκτιμητής Διαφορών σε Διαφορές

Από το συνδυασμό των δύο προηγούμενων μεθόδων, προκύπτει μία νέα μέθοδος αυτή των διαφορών σε διαφορές (*difference in differences*). Με τη διαστρωματική σύγκριση των ομάδων αποφεύγεται το πρόβλημα των χρονικών τάσεων συγκρίνοντας τις δύο ομάδες την ίδια χρονική περίοδο. Από την άλλη μεριά, με τη χρονολογική σύγκριση αποφεύγεται το πρόβλημα των απαρατήρητων συστηματικών διαφορών που μπορεί να έχουν δύο ομάδες, συγκρίνοντας τις ίδιες διαστρωματικές μονάδες της ομάδας αντιμετώπισης πριν και μετά από την απότομη μεταβολή. Ο εκτιμητής των διαφορών σε διαφορές (*DD estimator*) συνδυάζει αυτούς τους δύο εκτιμητές έτσι ώστε να επωφεληθεί από τα πλεονεκτήματα του καθενός.

Για την εφαρμογή της μεθόδου των διαφορών σε διαφορές, απαιτείται ένα σύνολο δεδομένων δύο περιόδων, μίας πριν και μίας μετά από το φυσικό πείραμα, καθώς και δύο ομάδες: η ομάδα ελέγχου και η ομάδα αντιμετώπισης. Το υπόδειγμα παλινδρόμησης της μεθόδου διαφορών σε διαφορές είναι της μορφής:

$$y = \beta_0 + \beta_1 d + \beta_2 p + \beta_3 (d \times p) + u \quad (2.8)$$

όπου d η ψευδομεταβλητή που περιγράφει σε ποια ομάδα ανήκει η κάθε παρατήρηση και λαμβάνει την τιμή ένα αν η παρατήρηση ανήκει στην ομάδα αντιμετώπισης και μηδέν αν ανήκει στην ομάδα ελέγχου, p η ψευδομεταβλητή που περιγράφει σε ποια

περίοδο αναφέρεται η παρατήρηση και λαμβάνει την τιμή ένα αν η παρατήρηση αναφέρεται στην περίοδο μετά τη μεταβολή ενώ την τιμή μηδέν αν αναφέρεται στην περίοδο προ της μεταβολής, ενώ η μεταβλητή ($d \times p$) συμβολίζει την αλληλεπίδραση των ψευδομεταβλητών p και d . Η ερμηνεία των συντελεστών β_0 , β_1 , β_2 και β_3 είναι η ακόλουθη: Ο συντελεστής β_0 αποτελεί το σταθερό όρο της εξίσωσης παλινδρόμησης, στον β_1 εμφανίζονται οι τυχόν συστηματικές διαφοροποιήσεις των ομάδων ελέγχου και αντιμετώπισης ενώ στον β_2 εμφανίζονται οι διαχρονικές τάσεις των ομάδων. Τέλος, ο συντελεστής β_3 εμφανίζει την πραγματική επίδραση της μεταβολής της πολιτικής στην ομάδα αντιμετώπισης. Είναι συνεπώς ο εκτιμητής των διαφορών σε διαφορές.

Οι αναμενόμενες τιμές της μεταβλητής απόκρισης της εξίσωσης (2.8) για κάθε συνδυασμό περιόδων και ομάδων δίνονται στη συνέχεια:

$$E(y|d = 1, p = 1) = \beta_0 + \beta_1 + \beta_2 + \beta_3,$$

$$E(y|d = 1, p = 0) = \beta_0 + \beta_1,$$

$$E(y|d = 0, p = 1) = \beta_0 + \beta_2,$$

$$E(y|d = 0, p = 0) = \beta_0,$$

υποθέτοντας ότι η αναμενόμενη τιμή του σφάλματος είναι ίση με μηδέν, δηλαδή $E(u|d, p) = 0$. Ταξινομώντας τις αναμενόμενες τιμές σε έναν πίνακα 2×2 και υπολογίζοντας τις διαφορές των στηλών και των γραμμών προκύπτει ο Πίνακας 2.2:

Πίνακας 2.2

Επιδράσεις πολιτικής ανά ομάδα και ανά περίοδο

Ο εκτιμητής των Διαφορών σε Διαφορές

	Μετά Μεταβολής	Προ Μεταβολής	Διαφορά
Αντιμετώπισης	$\beta_0 + \beta_1 + \beta_2 + \beta_3$	$\beta_0 + \beta_1$	$\beta_2 + \beta_3$
Ελέγχου	$\beta_0 + \beta_2$	β_0	β_2
Διαφορά	$\beta_1 + \beta_3$	β_1	β_3

Οι γραμμές «Αντιμετώπισης» και «Ελέγχου» καθώς και οι στήλες «Μετά Μεταβολής» και «Προ Μεταβολής» του πίνακα εμπεριέχουν τις αναμενόμενες τιμές για όλους τους συνδυασμούς περιόδου και ομάδας. Επομένως, η αναμενόμενη τιμή των διαστρωματικών μονάδων που ανήκουν στην ομάδα αντιμετώπισης για τη χρονική περίοδο μετά τη μεταβολή της πολιτικής, είναι $(\beta_0 + \beta_1 + \beta_2 + \beta_3)$ ενώ για την περίοδο πριν από τη μεταβολή είναι $(\beta_0 + \beta_1)$. Ομοίως η αναμενόμενη τιμή της

μεταβλητής απόκρισης για τις μονάδες που ανήκουν στην ομάδα ελέγχου για την περίοδο μετά τη μεταβολή είναι $(\beta_0 + \beta_2)$, ενώ πριν τη μεταβολή είναι (β_0) . Από την άλλη μεριά η στήλη και η γραμμή των διαφορών, εμπεριέχουν τις διαφορές των αναμενόμενων τιμών ανά ομάδα και ανά περίοδο αντίστοιχα. Ως εκ τούτου, η μέση διαφορά της μεταβλητής απόκρισης ανάμεσα στις δύο ομάδες για την περίοδο πριν από τη μεταβολή είναι ίση με $(\beta_0 + \beta_1) - \beta_0 = \beta_1$. Διαισθητικά γίνεται αντιληπτό ότι οι οριζόντιες διαφορές του πίνακα απαλείφουν την όποια χρονολογική τάση για κάθε ομάδα, ενώ οι κάθετες διαφορές απαλείφουν την όποια συστηματική διαφορά μεταξύ των ομάδων.

Ο εκτιμητής β_3 που αντιστοιχεί στη στήλη και στη γραμμή των διαφορών είναι αυτός των διαφορών σε διαφορές. Από τον τρόπο υπολογισμού του γίνεται αντιληπτό ότι είναι απαλλαγμένος από χρονολογικές τάσεις και από συστηματικές διαφορές. Υπολογίζεται είτε αφαιρώντας τους συντελεστές της στήλης των «Διαφορών», είτε τους συντελεστές της γραμμής των «Διαφορών». Στην πράξη ο εκτιμητής διαφορών σε διαφορές προκύπτει είτε αφαιρώντας τις συστηματικές διαφορές των ήδη απαλλαγμένων από χρονολογικές τάσεις ομάδων, είτε αφαιρώντας τις χρονολογικές τάσεις των ήδη απαλλαγμένων από συστηματικές διαφορές περιόδων. Αυτές οι δύο περιπτώσεις είναι ισοδύναμες και εκφράζονται από τις παρακάτω σχέσεις:

$$\beta_3 = [E(y|d = 1, p = 1) - E(y|d = 1, p = 0)] - [E(y|d = 0, p = 1) - E(y|d = 0, p = 0)]$$

ή

$$\beta_3 = [E(y|d = 1, p = 1) - E(y|d = 0, p = 1)] - [E(y|d = 1, p = 0) - E(y|d = 0, p = 0)]$$

Όπως σε κάθε ανάλυση εξίσωσης παλινδρόμησης, έτσι και εδώ, πέρα από την εύρεση του εκτιμητή των διαφορών σε διαφορές, υπάρχει ενδιαφέρον και για τη σημαντικότητά του. Επομένως, είναι απαραίτητη η εφαρμογή ενός στατιστικού ελέγχου της μορφής:

$$H_0: \beta_3 = 0$$

$$H_1: \beta_3 \neq 0$$

Ο έλεγχος της στατιστικής σημαντικότητας γίνεται με την εύρεση της στατιστική t , του συντελεστή β_3 , δηλαδή:

$$t = \frac{\hat{\beta}_3}{s(\hat{\beta}_3)}$$

όπου $\hat{\beta}_3$ η εκτίμηση του συντελεστή β_3 και $s(\hat{\beta}_3)$ το τυπικό του σφάλμα. Ο συντελεστής β_3 είναι στατιστικά σημαντικός αν $|t| > |t_{n-k-1, \alpha/2}|$, όπου α το επιθυμητό επίπεδο σημαντικότητας, n το πλήθος των παρατηρήσεων και k το πλήθος των ανεξάρτητων μεταβλητών (στο υπόδειγμα των διαφορών σε διαφορές $k = 3$).

2.5 Αμεροληψία Εκτιμητών Απλών Διαφορών

Η αξιοπιστία των αποτελεσμάτων ενός εκτιμητή εξαρτάται από την αμεροληψία του που εκφράζεται μαθηματικά ως:

$$E(\hat{\beta}_i) = \beta_i \quad (2.9)$$

Η εξέταση της αμεροληψίας των απλών εκτιμητών της παραγράφου (2.2) γίνεται μέσα από τις σχέσεις (2.8) και (2.9).

Σύμφωνα με τη σχέση (2.8) ο εκτιμητής που προκύπτει από τη σύγκριση των δύο ομάδων για την περίοδο μετά τη μεταβολή θα είναι ίσος με:

$$E(y|d = 1, p = 1) - E(y|d = 0, p = 1) = (\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_2)$$

άρα

$$E(y|d = 1, p = 1) - E(y|d = 0, p = 1) = \beta_1 + \beta_3$$

Ο εκτιμητής αυτός θα είναι αμερόληπτος εκτιμητής της επίδρασης της νέας πολιτικής, μόνο αν $\beta_1 = 0$. Στην πράξη για να χρησιμοποιηθεί αυτή η μέθοδος σαν ασφαλής μέθοδος εκτίμησης, θα πρέπει η ομάδα αντιμετώπισης και η ομάδα ελέγχου να μην εμφανίζουν συστηματική διαφορά, πράγμα αδύνατο στα περισσότερα φυσικά πειράματα.

Αντίστοιχα ο εκτιμητής που προκύπτει από τη σύγκριση της ομάδας αντιμετώπισης πριν και μετά από την εφαρμογή της νέας πολιτικής, σύμφωνα με τη σχέση (2.8) είναι:

$$E(y|d = 1, p = 1) - E(y|d = 1, p = 0) = (\beta_0 + \beta_1 + \beta_2 + \beta_3) - (\beta_0 + \beta_1)$$

άρα

$$E(y|d = 1, p = 1) - E(y|d = 1, p = 0) = \beta_2 + \beta_3$$

Ο εκτιμητής αυτός θα είναι αμερόληπτος εκτιμητής της επίδρασης της νέας πολιτικής, μόνο αν $\beta_2 = 0$. Στην πράξη για να χρησιμοποιηθεί αυτή η μέθοδος σαν ασφαλής μέθοδος εκτίμησης, θα πρέπει να μην υπάρχουν διαχρονικές τάσεις που επηρεάζουν τη μεταβλητή απόκρισης.

Ο εκτιμητής των διαφορών σε διαφορές μεριμνά για αυτές τις δύο αδυναμίες που αναφέρθηκαν προηγουμένως. Οποιαδήποτε συστηματική διαφορά μεταξύ των ομάδων

αντιμετώπισης και ελέγχου υπάρχει, αντιμετωπίζεται με την προσθήκη της ψευδομεταβλητής d στη (2.8), ενώ οποιαδήποτε χρονική τάση υπάρχει αντιμετωπίζεται με την προσθήκη της μεταβλητής p .

2.6 Υποθέσεις για το Υπόδειγμα Διαφορών σε Διαφορές

Το σύνολο των υποθέσεων που ισχύουν για το υπόδειγμα ελαχίστων τετραγώνων ισχύουν και για την εκτίμηση των συντελεστών σε ένα υπόδειγμα διαφορών σε διαφορές. Παρόλα αυτά υπάρχει άλλη μία προϋπόθεση από την οποία εξαρτάται άμεσα η αποτελεσματικότητα των εκτιμητών, αυτή της παράλληλης τάσης (*parallel trend*). Η προϋπόθεση αυτή δίνεται στη συνέχεια και αποτελείται από τρεις σχέσεις.

$$Cov(u, d) = 0$$

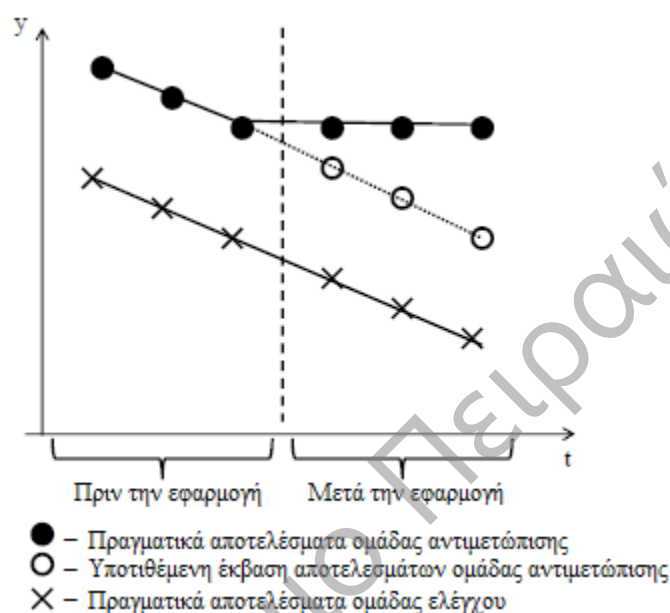
$$Cov(u, p) = 0$$

$$Cov(u, d \times p) = 0$$

Η πρώτη σχέση δηλώνει ότι τα σφάλματα του υποδείματος είναι ασυσχέτιστα με την ψευδομεταβλητή d , η δεύτερη δηλώνει ότι τα σφάλματα είναι ασυσχέτιστα με την ψευδομεταβλητή p , ενώ η τρίτη ότι τα σφάλματα είναι ασυσχέτιστα με την αλληλεπίδραση $d \times p$. Στην ουσία εφόσον οι σχέσεις αυτές κάνουν λόγο για σφάλματα ανεξάρτητα των ερμηνευτικών μεταβλητών, αναφέρονται στη μη ύπαρξη ενδογένειας. Η προϋπόθεση της παράλληλης τάσης αποτελεί την πιο σημαντική από το σύνολο των προϋποθέσεων, μιας και η βασική παραδοχή της μεθόδου των διαφορών σε διαφορές είναι ότι αν δεν υπάρχει μεταβολή στην πολιτική, η μέση μεταβολή της μεταβλητής απόκρισης θα είναι ίδια τόσο για την ομάδα αντιμετώπισης όσο και για την ομάδα ελέγχου. Στα διαγράμματα που ακολουθούν παρουσιάζεται η περίπτωση ύπαρξης παράλληλης τάσης καθώς και η περίπτωση μη ύπαρξης με τα προβλήματα που δημιουργεί.

Στο Διάγραμμα 2.1 απεικονίζεται η ύπαρξη της παράλληλης τάσης με το σχεδιασμό της αναμενόμενης απόκρισης των ομάδων αντιμετώπισης και ελέγχου για τις περιόδους πριν και μετά από το εφαρμογή της νέας πολιτικής. Για να τονιστεί η έννοια των παράλληλων τάσεων, υποθέτουμε ότι υπάρχουν τρεις διαφορετικές εκβάσεις για τις περιόδους. Τα αποτελέσματα των ομάδων αντιμετώπισης και ελέγχου συμβολίζονται με γεμάτους κύκλους και " \times " αντίστοιχα, στις συνεχόμενες γραμμές, ενώ με τους άδειους κύκλους στη διακεκομμένη γραμμή συμβολίζεται η έκβαση των

αποτελεσμάτων για την ομάδα αντιμετώπισης, αν δεν υπήρχε μεταβολή στην πολιτική. Η κάθετη διακεκομμένη γραμμή ορίζει τη στιγμή εφαρμογής της νέας πολιτικής και οι περίοδοι πριν από την εφαρμογή είναι τα σημεία του οριζόντιου άξονα αριστερά της διακεκομμένης ενώ οι περίοδοι μετά από την εφαρμογή τα σημεία του οριζόντιου άξονα δεξιά της διακεκομμένης.



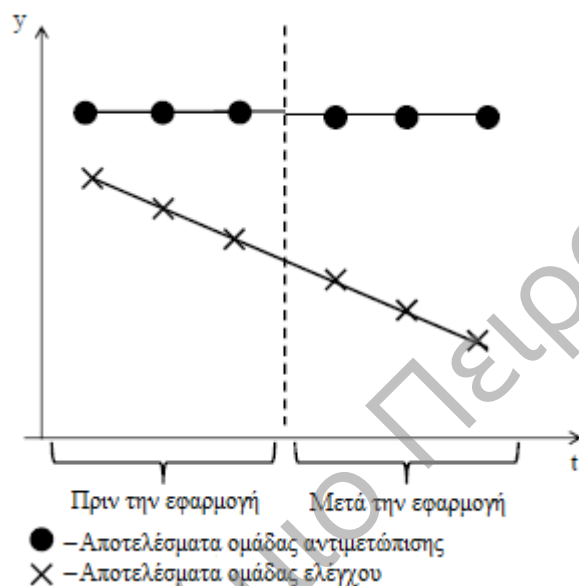
Διάγραμμα 2.1

Γραφική απεικόνιση της μεθόδου των διαφορών σε διαφορές

Στο Διάγραμμα 2.1 τα αποτελέσματα των δύο ομάδων διαφέρουν. Πιο συγκεκριμένα τα αποτελέσματα της ομάδας αντιμετώπισης είναι υψηλότερα από αυτά της ομάδας ελέγχου και έχουν πτωτική τάση με τον ίδιο ρυθμό μεταβολής, κατά τη διάρκεια των περιόδων, πριν από την εφαρμογή του συμβάντος. Αυτά τα δύο φαινόμενα, όπως αναφέρθηκε στην παράγραφο 2.4, δεν πρέπει δημιουργούν ανησυχία μιας και λαμβάνονται υπόψη στη μέθοδο των διαφορών σε διαφορές με την προσθήκη των ψευδομεταβλητών d και p , αντίστοιχα. Τη χρονική στιγμή της μεταβολής της πολιτικής παρατηρείται στα αποτελέσματα της ομάδας αντιμετώπισης, μία απότομη αλλαγή στην κλίση της τάσης, ενώ τα αποτελέσματα της ομάδας ελέγχου εξακολουθούν να έχουν τον ίδιο ρυθμό μεταβολής.

Το παραπάνω διάγραμμα τονίζει τη σημαντικότητα των παράλληλων τάσεων. Ενώ οι συστηματικές διαφορές αλλά και οι χρονολογικές τάσεις στις δύο ομάδες αντιμετωπίζονται επιτυχώς από τον εκτιμητή των διαφορών σε διαφορές, διαφορετικές

τάσεις μεταξύ των δύο ομάδων οδηγούν σε εσφαλμένα αποτελέσματα και σε μη αμερόληπτους εκτιμητές διαφορών σε διαφορές. Στο Διάγραμμα 2.2 παρουσιάζεται αυτό το πρόβλημα. Παρόλο που δεν υπάρχει καμία μεταβολή στη συμπεριφορά των δύο ομάδων μετά τη μεταβολή της πολιτικής, ο εκτιμητής των διαφορών σε διαφορές θα εκτιμήσει μία μεγάλη θετική επίδραση απλώς και μόνο λόγω της ύπαρξης διαφορετικών τάσεων. Επομένως, το πρόβλημα στην περίπτωση αυτή είναι ότι η επίδραση της εφαρμογής ενός νέου νόμου δε μπορεί να προσδιοριστεί.



Διάγραμμα 2.2

Παραβίαση της υπόθεσης των παράλληλων τάσεων

Όπως συμβαίνει με όλα τα προβλήματα ενδογένειας, δεν υπάρχει σαφής τρόπος να ελεγχθεί η υπόθεση των παράλληλων τάσεων και έτσι ακολουθούνται εμπειρικές μέθοδοι και η διαίσθησή του εκάστοτε ερευνητή. Ένας τρόπος ελέγχου της ύπαρξης παράλληλης τάσης είναι η συλλογή δεδομένων περισσότερων περιόδων πριν από την παρέμβαση και η οπτική σύγκριση των τάσεων των ομάδων αντιμετώπισης και ελέγχου για τις περιόδους αυτές. Ίσως ακόμη να απαιτείται και η αναζήτηση άλλων ομάδων ελέγχου που να παρέχουν διαφορετικές τάσεις. Στο Διάγραμμα 2.1 φαίνεται καθαρά η ύπαρξη παράλληλης τάσης ενώ στο Διάγραμμα 2.2 είναι οφθαλμοφανές ότι δεν ικανοποιείται η συνθήκη αυτή. Παρόλα αυτά, η ύπαρξη παράλληλων τάσεων δε συνεπάγεται και επίλυση του προβλήματος της ενδογένειας. Εάν υπάρχουν χρονικά μεταβαλλόμενες μεταβλητές που δεν έχουν συμπεριληφθεί στο υπόδειγμα και σχετίζονται με τη μεταβλητή απόκρισης, οι εκτιμητές των συντελεστών θα είναι ασυνεπείς.

2.7 Προεκτάσεις του Υποδείγματος Διαφορών σε Διαφορές

Η εκτίμηση και η αξιολόγηση της επίδρασης μίας νέας πολιτικής στην ομάδα αντιμετώπισης, μπορεί να πραγματοποιηθεί πέρα από την εξίσωση (2.8) και μέσω κάποιων παραλλαγών αυτής. Πιο συγκεκριμένα πέρα από τις ψευδομεταβλητές d , p και την αλληλεπίδρασή τους $d \times p$, είναι δυνατό να χρησιμοποιηθεί και πλήθος k ερμηνευτικών μεταβλητών που θεωρείται πως επηρεάζουν τη μεταβλητή απόκρισης, καθώς και η μεθοδολογία της απαλοιφής των απαρατήρητων επιδράσεων σε περίπτωση που το σύνολο δεδομένων έχει τη μορφή πάνελ. Οι δύο αυτές προεκτάσεις του υποδείγματος των διαφορών σε διαφορές, παρουσιάζονται αναλυτικά στη συνέχεια.

Εισαγωγή πλήθους k ερμηνευτικών μεταβλητών

Στην εξίσωση (2.8) μπορεί να συμπεριληφθεί πλήθος μεταβλητών X που θεωρείται ότι επηρεάζουν τη μεταβλητή απόκρισης. Το υπόδειγμα συνεπώς είναι της μορφής:

$$y = \beta_0 + \beta_1 d + \beta_2 p + \beta_3 (d \times p) + X\Gamma + u$$

όπου X το διάνυσμα ενός πλήθους k ερμηνευτικών μεταβλητών και Γ το διάνυσμα των συντελεστών των μεταβλητών αυτών. Πιο συγκεκριμένα $X = (x_1, \dots, x_k)$ και $\Gamma = (\gamma_1, \dots, \gamma_k)$. Σε κάθε περίπτωση ο συντελεστής που εκφράζει την όποια επίδραση από την έλευση ενός σημαντικού γεγονότος, είναι ο β_3 όπως και στην περίπτωση χωρίς τις k ερμηνευτικές μεταβλητές. Αν οι ομάδες ελέγχου και αντιμετώπισης έχουν επιλεγεί τυχαία, η εκτίμηση ελαχίστων τετραγώνων του β_3 είναι πιο αποτελεσματική με την προσθήκη και άλλων ερμηνευτικών μεταβλητών επειδή μειώνεται η διακύμανση του σφάλματος. Είναι σημαντικό να τονιστεί πως οι ερμηνευτικές αυτές μεταβλητές χρησιμοποιούνται για υποστηρικτικό σκοπό και πρέπει να μην επηρεάζονται από την εφαρμογή της πολιτικής.

Γενικεύοντας περαιτέρω το υπόδειγμα (2.8) ο εκάστοτε ερευνητής πέρα από τις μεταβλητές X μπορεί να συμπεριλάβει πολλαπλές χρονικές περιόδους και πολλαπλές ομάδες αντιμετώπισης. Το υπόδειγμα που έχουν προτείνει οι Bertrand, Duflo και Mullainathan (2003) είναι της μορφής:

$$y_{igt} = \beta_1 d_{gt} + X_{igt}\Gamma + p_t + m_g + u_{igt},$$

όπου οι δείκτες i, g και t αναφέρονται στις διαστρωματικές μονάδες, στις ομάδες και στις χρονικές περιόδους αντίστοιχα, με την d_{gt} να είναι μία ψευδομεταβλητή που λαμβάνει την τιμή ένα αν η ομάδα g τη χρονική περίοδο t έχει επηρεαστεί από την παρέμβαση, ενώ την τιμή μηδέν σε διαφορετική περίπτωση, το διάνυσμα X_{igt} να περιλαμβάνει το σύνολο των συνεχών ερμηνευτικών μεταβλητών, p_t οι σταθερές επιδράσεις της περιόδου, m_g οι σταθερές επιδράσεις της ομάδας και u_{igt} ο όρος σφάλματος. Η επίδραση της παρέμβασης δίνεται από την εκτίμηση του συντελεστή β_1 .

Ανάλυση Πολιτικής με Δεδομένα Πάνελ Δύο Περιόδων

Τα σύνολα δεδομένων πάνελ που αναφέρθηκαν στην Παράγραφο 1.5 είναι πολύ χρήσιμα για την ανάλυση πολιτικής και ειδικότερα για την αξιολόγηση προγράμματος. Στο απλούστερο πρόγραμμα αξιολόγησης, λαμβάνεται κατά την πρώτη περίοδο ένα δείγμα ατόμων, επιχειρήσεων, χωρών και ούτω καθεξής. Μερικές από αυτές τις μονάδες, αυτές που ανήκουν στην ομάδα αντιμετώπισης, συμμετέχουν σε ένα συγκεκριμένο πρόγραμμα σε μία μεταγενέστερη χρονική περίοδο, ενώ οι υπόλοιπες ανήκουν στην ομάδα ελέγχου. Η διαδικασία που περιγράφεται είναι πανομοιότυπη με αυτή των φυσικών πειραμάτων, με τη διαφορά όμως ότι σε αυτή την περίπτωση εμφανίζονται οι ίδιες διαστρωματικές μονάδες.

Έστω ότι η y_{it} συμβολίζει τη μεταβλητή απόκρισης, η ψευδομεταβλητή p_t λαμβάνει την τιμή ένα αν η παρατήρηση αφορά την περίοδο μετά την εφαρμογή του προγράμματος και μηδέν αν αφορά την περίοδο πριν από την εφαρμογή και η ψευδομεταβλητή d_{it} λαμβάνει την τιμή ένα αν η παρατήρηση έχει λάβει μέρος στο πρόγραμμα και μηδέν αν δεν έχει λάβει. Το απλούστερο υπόδειγμα παρατηρήτων επιδράσεων είναι της μορφής:

$$y_{it} = \beta_0 + \beta_1 p_t + \beta_2 d_{it} + \alpha_i + u_{it}$$

Λαμβάνοντας τις διαφορές για την απαλοιφή των παρατηρήτων επιδράσεων προκύπτει:

$$\Delta y_i = \beta_1 + \beta_2 \Delta d_i + \Delta u_{it} \quad (2.10)$$

Η επίδραση του προγράμματος στην ομάδα αντιμετώπισης είναι ίση με:

$$E(\Delta y | \Delta d_i = 1) = \beta_1 + \beta_2 \quad (2.11)$$

ενώ στην ομάδα ελέγχου είναι ίση με:

$$E(\Delta y | \Delta d_i = 0) = \beta_1 \quad (2.12)$$

Αφαιρώντας τις σχέσεις (2.11) και (2.12) προκύπτει η επίδραση της εφαρμογής του προγράμματος στην ομάδα αντιμετώπισης η οποία είναι ίση με το συντελεστή β_2 . Ως εκ τούτου:

$$E(\Delta y | \Delta d_i = 1) - E(\Delta y | \Delta d_i = 0) = \beta_2$$

Αυτή είναι μία εκδοχή του εκτιμητή των διαφορών σε διαφορές με χρήση δεδομένων πάνελ για δύο ομαδοποιημένες διαστρωματώσεις. Με τα δεδομένα πάνελ υπάρχει το πλεονέκτημα ότι με το σχηματισμό των διαφορών λαμβάνονται υπόψη οι απαραίτητες επιδράσεις για την κάθε διαστρωματική μονάδα.

2.8 Ελέγχοντας την Εσωτερική Εγκυρότητα

Όπως παρουσιάστηκε στην Παράγραφο 2.5 η σημαντικότερη υπόθεση της μεθόδου των διαφορών σε διαφορές, αυτή της παράλληλης τάσης, δε μπορεί να ελεγχθεί μέσα από στατιστικές διαδικασίες. Εναλλακτικά παρουσιάζεται στη συνέχεια μία σειρά διαδικασιών που μπορούν να γίνουν έτσι ώστε να εξασφαλισθεί η εσωτερική εγκυρότητα του εκτιμητή των διαφορών σε διαφορές.

- *Έλεγχος για την ορθή δομή των δεδομένων 1:* Για τον έλεγχο αυτόν επαναλαμβάνουμε την ανάλυση των διαφορών σε διαφορές για περιόδους πριν από την εφαρμογή τη νέας πολιτικής υποθέτοντας ότι η μεταβολή αυτή συνέβη σε προγενέστερο έτος. Η εκτιμώμενη επίδραση από τη μεταβολή αυτή (συντελεστής β_3), θα πρέπει να είναι στατιστικά ασήμαντη, ώστε να θεωρηθεί πως η παρατηρούμενη αλλαγή οφείλεται στη μεταβολή της πολιτικής και όχι σε κάποιον εναλλακτικό παράγοντα. Παρόμοια πρακτική ακολούθησε ο Almeida (2012) θέλοντας να εξετάσει την επενδυτική πολιτική των εταιριών που το χρέος τους έληγε αμέσως μετά την πιστωτική κρίση του 2007. Έλαβε σαν ομάδα ελέγχου εταιρίες με μικρό χρέος και σαν ομάδα αντιμετώπισης εταιρείες με μεγάλο χρέος, με ημερομηνίες ωρίμανσης αμέσως μετά την κρίση ενώ η απότομη μεταβολή ήταν η ίδια η κρίση. Μία απλοποιημένη μορφή του υποδείγματος που χρησιμοποιήθηκε είναι η παρακάτω:

$$y = \beta_0 + \beta_1 d + \beta_2 p + \beta_3 (d \times p) + u,$$

όπου η μεταβλητή y περιγράφει την επένδυση, η ψευδομεταβλητή d λαμβάνει την τιμή ένα αν η παρατήρηση ανήκει στην ομάδα ελέγχου και μηδέν αν ανήκει στην ομάδα αντιμετώπισης και η ψευδομεταβλητή p λαμβάνει την τιμή ένα αν

η παρατήρηση ανήκει στην περίοδο μετά την κρίση και μηδέν αν ανήκει στην περίοδο πριν από αυτήν. Θεωρώντας στη συνέχεια ότι η κρίση συνέβη το 2006 και το 2005 αντίστοιχα, βεβαιώθηκε ότι οι μεταβολές που παρατηρήθηκαν στην ομάδα αντιμετώπισης οφείλονταν καθαρά και μόνο στο φαινόμενο της κρίσης.

- *Έλεγχος για την ορθή δομή των δεδομένων 2:* Στην περίπτωση που έχουμε συμπεριλάβει στο υπόδειγμα συνεχείς ερμηνευτικές μεταβλητές, βεβαιωνόμαστε δεν επηρεάζονται από το συμβάν, εισάγοντας τες στη θέση της μεταβλητής απόκρισης y στο υπόδειγμα (2.8) και ελέγχοντας εκ νέου τις επιδράσεις της μεταβολής της πολιτικής σε αυτές.
- *Χρήση πολλαπλών ομάδων ελέγχου και αντιμετώπισης:* Με τη χρήση περισσότερων από μία ομάδες ελέγχου και αντιμετώπισης μειώνουμε το ενδεχόμενο μεροληψίας.
- *Δημιουργία ισορροπημένες ομάδων θεραπείας και ελέγχου:* Οι ομάδες θεραπείας και ελέγχου πρέπει να έχουν σχετικά ίδιο πλήθος παρατηρήσεων. Αν το πλήθος των παρατηρήσεων παρόλα αυτά δεν είναι ίδιο, το πρόβλημα αυτό μπορεί να αντιμετωπιστεί με την εισαγωγή επιπλέον ερμηνευτικών μεταβλητών.
- *Χρονική στιγμή μεταβολής της συμπεριφοράς της ομάδας αντιμετώπισης:* Η μεταβολή στη συμπεριφορά των παρατηρήσεων της ομάδας αντιμετώπισης, θα πρέπει να αρχίσει να παρατηρείται τη χρονική στιγμή γύρω από το συμβάν. Μετακινούμενοι πιο μακριά από αυτή τη χρονική στιγμή του συμβάντος τα αποτελέσματά μας επηρεάζονται και από άλλους παράγοντες με αποτέλεσμα να απειλείται η εγκυρότητα της μελέτης.
- *Αντιστροφή του Συμβάντος:* Αν η εφαρμογή μιας νέας πολιτικής προκαλέσει αλλαγή στη συμπεριφορά των παρατηρήσεων της ομάδας αντιμετώπισης τότε αν όλοι οι υπόλοιποι παράγοντες παραμένουν ίσοι, η αναίρεση της μεταβολής αυτής θα πρέπει να επαναφέρει την συμπεριφορά των παρατηρήσεων της ομάδας αντιμετώπισης στα προ μεταβολής επίπεδα.

Είναι φανερό πως οι παραπάνω έλεγχοι έχουν καθαρά εμπειρικό χαρακτήρα και η εφαρμογή και η ορθότητά τους στηρίζεται στην πείρα του ερευνητή καθώς και στο θεωρητικό του υπόβαθρο. Τέλος, είναι πιθανό μόνο κάποιοι από αυτούς να απαιτούνται σε κάθε εμπειρική μελέτη.

2.9 Ανακεφαλαίωση

Στο κεφάλαιο αυτό αρχικά έγινε αναφορά στα προβλήματα που μπορεί να αντιμετωπίσει ένας ερευνητής στην εσωτερική και εξωτερική εγκυρότητα μίας εμπειρικής έρευνας. Στη συνέχεια αναλύθηκαν απλές μέθοδοι εκτίμησης των επιδράσεων μίας παρέμβασης (πολιτικής, οικονομικής, κυβερνητικής κ.α.) σε δύο ομάδες πληθυσμού για μία χρονική περίοδο και σε μία ομάδα πληθυσμού για δύο χρονικές περιόδους με τα προβλήματα που αντιμετωπίζουν, για να φτάσουμε στη ζητούμενη μέθοδο των διαφορών σε διαφορές. Έγινε αναφορά στις προϋποθέσεις που απαιτούνται για να είναι οι εκτιμητές των διαφορών σε διαφορές αμερόληπτοι, δόθηκε ιδιαίτερη σημασία στην υπόθεση των παράλληλων τάσεων και τέλος αναφέρθηκαν απλές μέθοδοι για την αποφυγή προβλημάτων εσωτερικής εγκυρότητας στο υπόδειγμα των διαφορών σε διαφορές.

Πανεπιστήμιο Πειραιώς

ΚΕΦΑΛΑΙΟ 3

ΑΝΑΦΟΡΑ ΣΕ ΕΜΠΕΙΡΙΚΕΣ ΜΕΛΕΤΕΣ

3.1 Εισαγωγή

Η ιδέα της χρήσης της μεθόδου των διαφορών σε διαφορές στις εμπειρικές μελέτες είναι αρκετά παλιά. Στις αρχές του 20^{ου} αιώνα και πιο συγκεκριμένα το 1915, οι Marie Louise Obenauer και Bertha Marie von der Nienburg χρησιμοποιούν τη μέθοδο αυτή, για να εξετάσουν τις επιπτώσεις της θεσμοθέτησης ενός κατώτατου εβδομαδιαίου μισθού στη γυναικεία απασχόληση, χρησιμοποιώντας σαν ομάδα ελέγχου άνδρες εργαζόμενους, που δεν επηρεάζονται από την όποια μεταβολή. Η μελέτη αυτή δημοσιεύεται σε έκθεση της εποχής από τη στατιστική αρχή (*Bureau of Labor Statistics - BLS*). Στις μέρες μας, πολλοί ερευνητές έχουν κάνει χρήση της μεθόδου των διαφορών σε διαφορές στις εμπειρικές τους μελέτες, με πιο γνωστή αυτή των David Card και Alan B. Krueger (1994), οι οποίοι μελετώντας τις επιπτώσεις της αύξησης του ωρομίσθιου, από τα \$4.25 στα \$5.05 στην πολιτεία του New Jersey, στους δείκτες της απασχόλησης, καταλήγουν σε εντελώς αντίθετα συμπεράσματα από αυτά των συμβατικών μελετών. Πιο συγκεκριμένα εξάγουν το συμπέρασμα ότι η αύξηση του ωρομίσθιου, οδηγεί σε αύξηση της απασχόλησης, ενώ μέχρι τότε οι κλασικές οικονομετρικές μελέτες αναφέρονται σε ακριβώς αντίθετα αποτελέσματα με μείωση της απασχόλησης.

Στον παρόν κεφάλαιο γίνεται ανασκόπηση στις μελέτες των Obenauer και Nienburg (1915), των David Card και Alan B. Krueger (1994), καθώς και στην μελέτη των Bruce D. Meyer, W. Kip Viscusi και David L. Durbin (1995), οι οποίοι μελετούν την επίδραση της αύξησης της αποζημίωσης στη διάρκεια της αναρρωτικής άδειας, μετά από εργατικό ατύχημα.

3.2 Κατώτατος Μισθός και Απασχόληση στις Γυναίκες του Oregon

Σύμφωνα με μία πράξη του 1913, με τη δημιουργία μίας Βιομηχανικής Επιτροπής Πρόνοιας για την προστασία της ζωής, της υγείας και της ηθικής των γυναικών και των

νέων ηλικιακά εργαζόμενων γυναικών, καθιερώθηκε κατώτατος εβδομαδιαίος μισθός ύψους \$9.25 για εργαζόμενες στο Portland, με περισσότερο από ένα χρόνο εμπειρίας στην αντίστοιχη θέση. Για τις γυναίκες χωρίς επαγγελματική εμπειρία αλλά και για κορίτσια ηλικίας 16 έως 18 ετών, καθιερώθηκε κατώτατος εβδομαδιαίος μισθός ύψους \$6, ενώ παράλληλα ορίστηκε ως μέγιστος εβδομαδιαίος χρόνος εργασίας αυτός των 50 ωρών. Μία από τις πρώτες εμπειρικές μελέτες για τις επιδράσεις του κατώτατου ωρομίσθιου στην απασχόληση έλαβε χώρα από τις Obenauer και Nienburg το 1915. Χρησιμοποίησαν δεδομένα μισθοδοσίας και προσωπικών συνεντεύξεων για γυναίκες και άντρες σε καταστήματα λιανικής στο Oregon των Ηνωμένων Πολιτειών, πριν και μετά την εφαρμογή κατώτατου ωρομίσθιου στις γυναίκες εργαζόμενες για τα έτη 1913 και 1914. Τα δεδομένα των γυναικών αφορούσαν την ομάδα αντιμετώπισης, ενώ τα δεδομένα των ανδρών την ομάδα ελέγχου. Οι Obenauer και Nienburg συνέλλεξαν δεδομένα απευθείας από τα βιβλία μισθοδοσίας 30 καταστημάτων λιανικής πώλησης (και επτά ακόμη από το Salem, όπου ο κατώτατος εβδομαδιαίος μισθός για τις γυναίκες με προϋπηρεσία ισοδυναμούσε με \$8.25). Τα δεδομένα του Portland κάλυψαν 1546 γυναίκες και 868 άνδρες, με πληροφορίες σχετικά με τα ποσοστά των μισθών και των ωρών εργασίας για τις γυναίκες και τα συνολικά κέρδη για τους άνδρες. Τα στοιχεία αυτά, συμπληρώθηκαν από ατομικές συνεντεύξεις 443 γυναικών, προκειμένου να ληφθούν λεπτομερείς πληροφορίες σχετικά με τις αλλαγές στο καθεστώς της απασχόλησης που θα μπορούσε να αποδοθεί στην εισαγωγή του κατώτατου μισθού.

Οι κύριες επιπτώσεις στην απασχόληση στα δεδομένα Όρεγκον παρατίθενται στον Πίνακα 3.1 ο οποίος αποτελεί μία καλή εισαγωγή στα εμπειρικά αποτελέσματα των Card and Krueger, διότι η μελέτη βασίστηκε στην ιδέα ενός «φυσικού πειράματος» και επειδή η αύξηση του κατώτατου μισθού οδήγησε σε αύξηση της εργασίας σε ορισμένες περιπτώσεις.

Πίνακας 3.1

Κύριες επιπτώσεις στην απασχόληση

	Men	Girls (Age 16-18)	Ratio (Girls/Men)	Women (Age > 18)	Ratio (Women/Men)	Women Age Unknown
Before (Mar/Ap 1913)	940	138	.1468	1543	1.641	152
After Mar/Ap 1914)	868	160	.1843	1327	1.529	59
Change	-72	22	.0375	-216	-0.113	-93
% Change	-7.7%	15.9%	23.6%	-14%	-6.3%	-61.2%

Πηγή: Obenauer και Nienburg, 1915

Οι μετρήσεις της απασχόλησης για τους άνδρες και τις γυναίκες δεν ήταν ανάλογες, και έτσι αντί να χρησιμοποιηθεί η μέθοδος των διαφορών σε διαφορές για τα επίπεδα απασχόλησης, χρησιμοποιήθηκε η μεταβολή στους δείκτες της απασχόλησης, με στόχο την ποσοτικοποίηση των επιπτώσεων της θεσμοθέτησης κατώτατου μισθού. Με τη μέτρηση αυτή, η ενήλικη γυναικεία εργασία μειώθηκε κατά 6.3% ενώ η νεανική γυναικεία εργασία αυξήθηκε κατά 23.6%. Οι Obenauer και Nienburg επεσήμαναν ότι η συνολική μείωση της απασχόλησης οφείλεται σε μία γενικότερη ύφεση: για παράδειγμα, οι συνολικές πωλήσεις στα καταστήματα του Portland μειώθηκαν κατά 8.6% κατά την περίοδο αυτή. Επιπλέον σημείωσαν ότι οι θέσεις εργασίας των ανδρών ήταν λιγότερο ευάλωτες σε αυτές τις μειώσεις προσωπικού από αυτές των γυναικών, με αποτέλεσμα η εκτίμηση μέσω της μεθόδου των διαφορών σε διαφορές, να μη συλλαμβάνει την όποια επίδραση από τη μεταβολή του κατώτατου μισθού. Κατέληξαν στο συμπέρασμα ότι οι μεταβολές στη γυναικεία απασχόληση, ελάχιστα οφείλονται στη μεταβολή του κατώτατου μισθού. Παρ' όλα αυτά υπήρξε ολοφάνερη απόδειξη, η οποία επιβεβαιώθηκε από τις συνεντεύξεις, ότι οι γυναίκες με επαγγελματική εμπειρία στις θέσεις χαμηλών προσόντων, έχασαν τις δουλειές τους και αντικαταστάθηκαν από νεοεισερχόμενες στην αγορά εργασίας αλλά και μαθητευόμενες, οι οποίες πληρώνονταν με \$6 ανά εβδομάδα έναντι \$9.25 που απαιτούνταν για τις πιο έμπειρες.

Ένα άλλο ενδιαφέρον χαρακτηριστικό της μελέτης των Obenauer και Nienburg εμφανίζεται στον Πίνακα 3.2 και αφορά τη λεπτομερή καταμέτρηση της κατανομής των μισθών πριν και μετά από την εφαρμογή του κατώτατου μισθού, για τις 374 γυναίκες που συμμετείχαν στην έρευνα και οι οποίες είχαν απασχοληθεί στο Oregon πριν από το 1914. Τα αποτελέσματα του Πίνακα 3.2 συγκεντρώνονται στο Διάγραμμα 3.1, όπου φανερώνεται ότι στις περισσότερες περιπτώσεις (204 από 374), δεν υπήρξε καμία μεταβολή των αμοιβών μεταξύ του 1913 και του 1914. Παρόλα αυτά υπήρξαν πολλές γυναίκες με μισθολογικές αυξήσεις, είτε στο νέο κατώτατο μισθό (73) είτε και σε πιο υψηλά επίπεδα (82), ενώ τέλος σε ελάχιστες περιπτώσεις ο μισθός μειώθηκε (15).

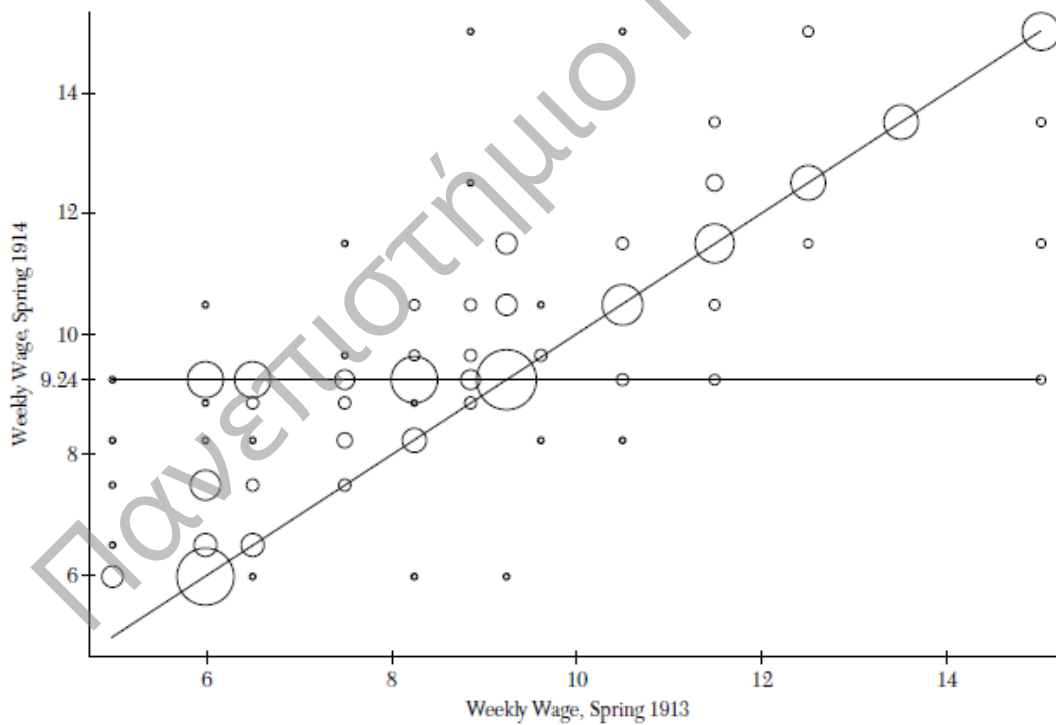
Πίνακας 3.2

Μισθοί πριν και μετά τη θεσμοθέτηση κατώτατου μισθού

Weekly rate of pay before minimum-wage determinations.	Weekly rate of pay after minimum-wage determinations.												Total number.	Per cent.
	\$6	\$6.01 to \$6.99	\$7 to \$7.99	\$8 to \$8.49	\$8.50 to \$9.22	\$9.23 to \$9.25	\$9.26 to \$9.99	\$10 to \$10.99	\$11 to \$11.99	\$12 to \$12.99	\$13 to \$13.99	\$14 and over.		
Under \$6.....	7	1	1	1	1	1							11	2.9
\$6.....	44	8	13	1	1	15		1					83	22.2
\$6.01 to \$6.99 ..	1	6	3	1	2	15							28	7.5
\$7 to \$7.99.....			3	4	2	5	1		1				16	4.3
\$8 to \$8.49.....	1			8	1	31	2	2					45	12.0
\$8.50 to \$9.22 ..					2	6	2	2		1			14	3.7
\$9.23 to \$9.25 ..	1					47	6	5					59	15.8
\$9.26 to \$9.99 ..				1			3	1					5	1.3
\$10 to \$10.99....				1		2		23	2			1	29	7.8
\$11 to \$11.99....						2		2	17	4	2		27	7.2
\$12 to \$12.99....									1	14			17	4.6
\$13 to \$13.99....											18		18	4.8
\$14 and over....						1			1		1		19	5.9
Total.....	54	15	20	17	8	125	8	37	27	19	21	23	374	100.0
Percent....	14.4	4.0	5.4	4.6	2.1	33.4	2.1	9.9	7.2	5.1	5.6	6.2	100.0

¹ Excludes 65 women who were not employed in Oregon prior to the minimum-wage determinations.
² Not including 4 women, whose previous rate was not reported; present rates, respectively, 2 at \$6, 1 at \$6.92, and 1 at \$14.

Πηγή: Obenauer και Nienburg, 1915



Πηγή: Obenauer και Nienburg, 1915

Διάγραμμα 3.1

Μισθολογική μεταβολή στο Oregon, Γυναίκες, 1913 – 1914

Το υπουργείο εργασίας των Ηνωμένων Πολιτειών Αμερικής, πραγματοποίησε παρόμοιες μελέτες μετά την εφαρμογή κατώτατου μισθού στις εργαζόμενες γυναίκες, στη Νέα Υόρκη και στο Ohio το 1934, χρησιμοποιώντας ως ομάδα αντιμετώπισης εργαζόμενες σε πλυντήρια της Νέας Υόρκης και του Ohio και ως ομάδα ελέγχου εργαζόμενες σε αντίστοιχους κλάδους από την Pennsylvania και την Indiana αντίστοιχα. Οι πρώτες μελέτες για την επίδραση της εισαγωγής ομοσπονδιακού κατώτατου μισθού το 1938, καθώς και μεταγενέστερες μελέτες για τις επιδράσεις της αύξησης αυτού κατά \$0.30 το 1939, \$0.75 το 1950 και \$1.00 το 1956 βασίστηκαν σε συγκρίσεις εργαζόμενων, που καλύπτονταν και δεν καλύπτονταν από τις μεταβολές ή σε συγκρίσεις χαμηλόμισθων και υψηλόμισθων εργαζόμενων.

Οι μελέτες αυτές, καθώς και αυτή των Obenauer και Nienburg, εξετάστηκαν σε μεταγενέστερο χρόνο από τον Peterson το 1959, ο οποίος αναγνώρισε ως γενικά αποδεκτό συμπέρασμα, ότι οι μεταβολές του κατώτατου μισθού δεν επηρεάζουν σημαντικά την απασχόληση, ενώ παράλληλα υποστήριξε πως οι όποιες μεταβολές ερμηνεύθηκαν ως σημαντικές, οφείλεται στην τάση των ερευνητών να ερμηνεύουν τα γεγονότα λανθασμένα. Κατά την άποψή του, μία πιο προσεκτική ανάγνωση των δεδομένων αποκαλύπτει τα αναμενόμενα αποτελέσματα με βάση την υπόθεση μιας αρνητικής κλίσης στην καμπύλη ζήτησης της εργασίας.

Ο Peterson, έκανε μία υπόθεση για τις αρνητικές επιπτώσεις στην απασχόληση, αλλά παραδέχθηκε ότι και άλλες λογικές ερμηνείες είναι δυνατές, κυρίως επειδή τα δεδομένα δεν προέρχονται από ένα ελεγχόμενο πείραμα. Ο Lester (1960), λαμβάνοντας υπόψη τα ίδια στοιχεία, κατέληξε στο συμπέρασμα ότι θα μπορούσε να συμβεί σχεδόν οτιδήποτε μετά από μία μέτρια αύξηση του κατώτατου μισθού.

3.3 Μεταβολή Κατώτατου Μισθού και Απασχόληση στο New Jersey

Την 1^η Απρίλη του 1992, το κατώτατο ωρομίσθιο στην πολιτεία του New Jersey (NJ) αυξήθηκε από τα \$4.25 στα \$5.05. Οι David Card και Alan B. Krueger, για να αξιολογήσουν τις επιπτώσεις αυτής της αύξησης στους δείκτες της απασχόλησης, συγκέντρωσαν δείγμα 410 ταχυφαγείων από τις πολιτείες του New Jersey και της ανατολικής Pennsylvania (PA) πριν και μετά από αυτή την αύξηση. Να σημειωθεί στο σημείο αυτό, ότι το κατώτατο ωρομίσθιο στην Pennsylvania παρέμεινε αμετάβλητο.

Συγκρίσεις στην αύξηση της απασχόλησης στα καταστήματα του New Jersey και της Pennsylvania παρείχαν απλές εκτιμήσεις της επίδρασης του υψηλότερου βασικού ωρομίσθιου, ενώ ακόμα σύγκριναν τη μεταβολή της απασχόλησης σε καταστήματα που η αμοιβή ήταν εξαρχής υψηλότερη των \$5 με αυτά που ήταν \$4.25 μόνο για την περιοχή του New Jersey και δεν εντόπισαν κάποια ένδειξη ότι η αύξηση του μισθού οδήγησε σε μείωση της απασχόλησης.

- **Ο Νόμος του New Jersey**

Το Νοέμβριο του 1989 υπογράφηκε ένα νομοσχέδιο που προέβλεπε αύξηση του κατώτατου ωρομίσθιου στο σύνολο των Η.Π.Α από τα \$3.35 στα \$3.80, με εφαρμογή την 1^η του Απρίλη του 1990 και με περαιτέρω αύξηση στα \$4.25 από την 1^η Απρίλη του 1991. Στις αρχές του 1990 το New Jersey προχώρησε ένα βήμα παραπέρα θεσπίζοντας πέραν των προκαθορισμένων αυξήσεων και νέα αύξηση στα \$5.05 με ημέρα εφαρμογής την 1^η του Απρίλη του 1992. Η αύξηση αυτή έχρησε το New Jersey πρωτοπόρο στο κατώτατο ωρομίσθιο, ξεσηκώνοντας παράλληλα αντιδράσεις στον επιχειρηματικό κόσμο.

Στα δύο χρόνια που μεσολάβησαν μέχρι την εφαρμογή του κατώτατου ωρομίσθιου στα \$5.05 η οικονομία του New Jersey μπήκε σε ύφεση. Υπό το φόβο αρνητικών επιπτώσεων, από την αύξηση αυτή το νομοθετικό σώμα ψήφισε το Μάρτιο του 1992 η αύξηση να γίνει σταδιακά αλλά τελικώς εφαρμόστηκε από την 1^η Απριλίου όπως και είχε προσυμφωνηθεί.

- **Σχεδιασμός του δείγματος και αξιολόγηση**

Στις αρχές του 1992 οι David Card και Alan B. Krueger, αποφάσισαν να αξιολογήσουν την ενδεχόμενη αύξηση, κάνοντας έρευνα σε ταχυφαγεία του New Jersey και της ανατολικής Pennsylvania. Η επιλογή των ταχυφαγείων έγινε για πολλούς λόγους. Πρώτον, τα ταχυφαγεία είναι πρωτοπόρα στην παροχή κατώτατων ωρομισθίων. Δεύτερον συμμορφώνονται με τη νομοθεσία και έτσι θα προχωρούσαν άμεσα σε αύξηση του κατώτατου ωρομισθίου. Τρίτον οι απαιτήσεις για τις θέσεις εργασίας καθώς και τα προϊόντα είναι σχετικά ομοιογενή, κάνοντας έτσι ευκολότερη τη λήψη αξιόπιστων μετρήσεων για την απασχόληση, τους μισθούς και τις τιμές των προϊόντων. Τέταρτον είναι πιο εύκολη η κατασκευή δείγματος από franchise

εστιατόρια και τέλος από προηγούμενη έρευνα των Katz και Krueger γνώριζαν ότι τα ταχυφαγεία απαντούν πιο εύκολα σε τηλεφωνικές έρευνες.

Βάσει των παραπάνω θεωρήσεων, ξεκίνησε η έρευνά στις αλυσίδες Burger King, KFC, Wendy's και Roy Rogers των δύο πολιτειών. Η πρώτη φάση της έρευνας έλαβε χώρα από τα τέλη του Φεβρουάριου του 1992 μέχρι και τις αρχές του Μάρτιου και συμπεριελάμβανε ερωτήσεις σχετικές με την απασχόληση, τους αρχικούς μισθούς και άλλες παρόμοιες ερωτήσεις ενώ η δεύτερη φάση της έρευνας έλαβε χώρα τους μήνες Νοέμβρη και Δεκέμβρη, οχτώ μήνες μετά από τη μεταβολή του κατώτατου μισθού. Τελικά υπήρξε μεγάλη ανταπόκριση όπως φαίνεται στον Πίνακα 3.3.

Πίνακας 3.3

Σχεδιασμός δείγματος και απώλειες

	All	Stores in:	
		NJ	PA
<i>Wave 1, February 15– March 4, 1992:</i>			
Number of stores in sample frame: ^a	473	364	109
Number of refusals:	63	33	30
Number interviewed:	410	331	79
Response rate (percentage):	86.7	90.9	72.5
<i>Wave 2, November 5– December 31, 1992:</i>			
Number of stores in sample frame:	410	331	79
Number closed:	6	5	1
Number under renovation:	2	2	0
Number temporarily closed: ^b	2	2	0
Number of refusals:	1	1	0
Number interviewed: ^c	399	321	78

^aStores with working phone numbers only; 29 stores in original sample frame had disconnected phone numbers.

^bIncludes one store closed because of highway construction and one store closed because of a fire.

^cIncludes 371 phone interviews and 28 personal interviews of stores that refused an initial request for a phone interview.

Πηγή: Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania, Card and Krueger, 1994

Στην πρώτη φάση ρωτήθηκαν 473 εστιατόρια, από τα οποία τα 364 στο New Jersey και τα 109 στην Pennsylvania. Από το σύνολο αυτό, αρνήθηκαν να λάβουν μέρος στην έρευνα 63, από τα οποία τα 33 στο New Jersey και τα 30 στην Pennsylvania. Στη δεύτερη φάση της έρευνας ρωτήθηκαν τα 410 που έλαβαν μέρος και στην πρώτη φάση, από τα οποία υπήρχε απώλεια 11 καταστημάτων συνολικά, 10 από την πρώτη πολιτεία και 1 από τη δεύτερη.

Στον Πίνακα 3.4 παρουσιάζονται οι μέσοι για αρκετές βασικές μεταβλητές του συνόλου των δεδομένων και παρουσιάζονται παράλληλα για τα εστιατόρια στο New

Jersey και την Pennsylvania, μαζί με τις στατιστικές t για την αρχική υπόθεση ότι οι μέσοι είναι ίδιοι στις δύο πολιτείες.

Στις πρώτες πέντε γραμμές παρουσιάζεται η κατανομή της κάθε αλυσίδας και το καθεστώς ιδιοκτησίας (franchise ή εταιρικό κατάστημα). Τα καταστήματα των Burger King, Roy Rogers και Wendy's στο δείγμα, έχουν ίδιες μέσες τιμές φαγητού, ώρες λειτουργίας και επίπεδα απασχόλησης. Τα καταστήματα KFC είναι μικρότερα και ανοίγουν για λιγότερες ώρες, ενώ προσφέρουν πιο ακριβά γεύματα από τις υπόλοιπες αλυσίδες.

Πίνακας 3.4

Κατανομή ταχυφαγείων και μέσοι όροι σημαντικών μεταβλητών

Variable	Stores in:		t^a
	NJ	PA	
1. <i>Distribution of Store Types (percentages):</i>			
a. Burger King	41.1	44.3	-0.5
b. KFC	20.5	15.2	1.2
c. Roy Rogers	24.8	21.5	0.6
d. Wendy's	13.6	19.0	-1.1
e. Company-owned	34.1	35.4	-0.2
2. <i>Means in Wave 1:</i>			
a. FTE employment	20.4 (0.51)	23.3 (1.35)	-2.0
b. Percentage full-time employees	32.8 (1.3)	35.0 (2.7)	-0.7
c. Starting wage	4.61 (0.02)	4.63 (0.04)	-0.4
d. Wage = \$4.25 (percentage)	30.5 (2.5)	32.9 (5.3)	-0.4
e. Price of full meal	3.35 (0.04)	3.04 (0.07)	4.0
f. Hours open (weekday)	14.4 (0.2)	14.5 (0.3)	-0.3
g. Recruiting bonus	23.6 (2.3)	29.1 (5.1)	-1.0
3. <i>Means in Wave 2:</i>			
a. FTE employment	21.0 (0.52)	21.2 (0.94)	-0.2
b. Percentage full-time employees	35.9 (1.4)	30.4 (2.8)	1.8
c. Starting wage	5.08 (0.01)	4.62 (0.04)	10.8
d. Wage = \$4.25 (percentage)	0.0	25.3 (4.9)	—
e. Wage = \$5.05 (percentage)	85.2 (2.0)	1.3 (1.3)	36.1
f. Price of full meal	3.41 (0.04)	3.03 (0.07)	5.0
g. Hours open (weekday)	14.4 (0.2)	14.7 (0.3)	-0.8
h. Recruiting bonus	20.3 (2.3)	23.4 (4.9)	-0.6

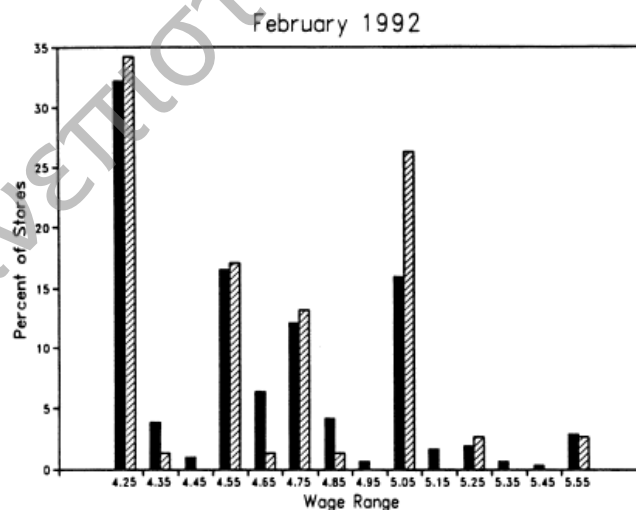
Notes: See text for definitions. Standard errors are given in parentheses.

^aTest of equality of means in New Jersey and Pennsylvania.

Πηγή: *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, Card and Krueger, 1994

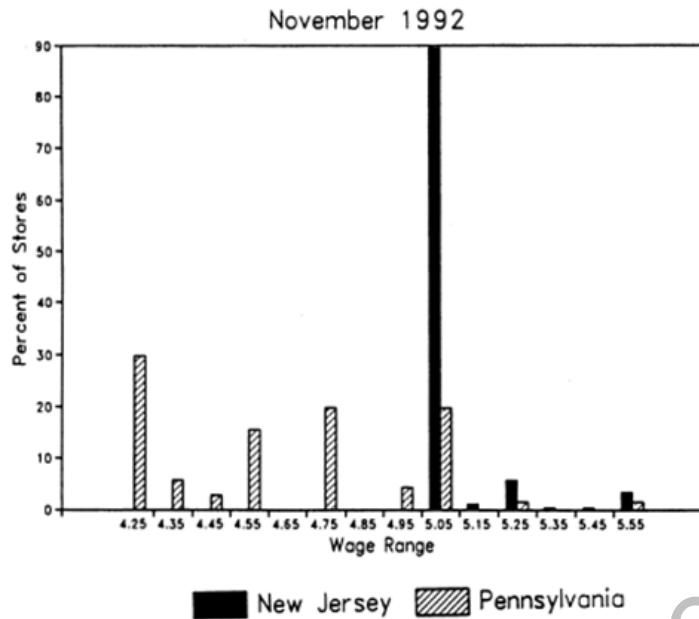
Στην πρώτη φάση της έρευνας, η μέση απασχόληση ήταν 23.3 εργαζόμενοι πλήρους απασχόλησης (*Full Time Equivalent - FTE*) σε κάθε κατάστημα στην Pennsylvania, έναντι 20.4 στο New Jersey. Οι αρχικοί μισθοί ήταν ίδιοι μεταξύ των καταστημάτων στις δύο πολιτείες, ωστόσο η τιμή ενός πλήρους γεύματος ήταν σημαντικά υψηλότερη στο New Jersey. Δεν υπήρχαν σημαντικές διαφορές στις ώρες εργασίας ανά εργαζόμενο πλήρους απασχόλησης, καθώς και στα πακέτα μπόνους.

Ο μέσος αρχικός μισθός στα ταχυφαγεία, αυξήθηκε κατά 10 ποσοστιαίες μονάδες εξαιτίας της αύξησης του κατώτατου ωρομισθίου. Παρά την αύξηση αυτή, η πλήρης απασχόληση αυξήθηκε στο New Jersey σε σχέση με την Pennsylvania. Λαμβάνοντας υπόψη ότι τα καταστήματα του New Jersey ήταν εξ' αρχής μικρότερα, η αύξηση της απασχόλησης σε συνδυασμό με τις απώλειες στην Pennsylvania οδήγησε σε μη σημαντικές διαφορές στη δεύτερη φάση της έρευνας. Μόνο δύο άλλες μεταβλητές παρουσίασαν μία σχετική αλλαγή μεταξύ των φάσεων 1 και 2. Οι εργαζόμενοι με πλήρη απασχόληση και η τιμή του βασικού γεύματος. Και οι δύο αυτές μεταβλητές αυξήθηκαν στο New Jersey σε σχέση με την Pennsylvania. Στα Διαγράμματα 3.2 και 3.3 παρουσιάζεται η κατανομή των αρχικών μισθών στις δύο πολιτείες, για την περίοδο πριν από την αύξηση του βασικού μισθού και μετά από αυτήν αντίστοιχα, όπου οι γραμμοσκιασμένες στήλες αναφέρονται στην Pennsylvania ενώ οι μαύρες στο New Jersey.



Διάγραμμα 3.2

Κατανομή κατώτατου ωρομισθίου πριν τη μεταβολή



Πηγή: *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, Card and Krueger, 1994

Διάγραμμα 3.3

Κατανομή κατώτατου ωρομίσθιου μετά τη μεταβολή

- **Επιδράσεις του κατώτατου ωρομίσθιου στην απασχόληση – Μέθοδος Διαφορών σε Διαφορές**

Στον Πίνακα 3.5 συγκεντρώνονται τα επίπεδα και οι μεταβολές της μέσης απασχόλησης πλήρους ωραρίου για κάθε κατάσταση της έρευνας. Στις στήλες (i) και (ii) παρουσιάζονται τα δεδομένα ανά πολιτεία, ενώ στις στήλες (iv), (v) και (vi) κατηγοριοποιούνται τα καταστήματα του New Jersey ανάλογα με το αν στην πρώτη φάση της έρευνας το ωρομίσθιο είναι ίσο με \$4.25, από \$4.26 έως \$4.99 και από \$5.00 και πάνω αντίστοιχα. Στη στήλη (iii) εμφανίζονται οι μεταβολές της μέσης πλήρους απασχόλησης, ανάμεσα στα καταστήματα του New Jersey και της Pennsylvania, ενώ στις στήλες (vii) και (viii), οι μεταβολές της μέσης απασχόλησης στις διάφορες κατηγορίες μισθών στο New Jersey και τέλος στην τρίτη γραμμή του Πίνακα 3.5 εμφανίζονται οι μεταβολές στη μέση απασχόληση πλήρους ωραρίου ανάμεσα στις δύο φάσεις της έρευνας.

Πίνακας 3.5

Μέση απασχόληση για κάθε κατάστημα πριν και μετά την αύξηση

Variable	Stores by state			Stores in New Jersey ^a			Differences within NJ ^b	
	PA (i)	NJ (ii)	Difference, NJ – PA (iii)	Wage = \$4.25 (iv)	Wage = \$4.26–\$4.99 (v)	Wage ≥ \$5.00 (vi)	Low– high (vii)	Midrange– high (viii)
1. FTE employment before, all available observations	23.33 (1.35)	20.44 (0.51)	–2.89 (1.44)	19.56 (0.77)	20.08 (0.84)	22.25 (1.14)	–2.69 (1.37)	–2.17 (1.41)
2. FTE employment after, all available observations	21.17 (0.94)	21.03 (0.52)	–0.14 (1.07)	20.88 (1.01)	20.96 (0.76)	20.21 (1.03)	0.67 (1.44)	0.75 (1.27)
3. Change in mean FTE employment	–2.16 (1.25)	0.59 (0.54)	2.76 (1.36)	1.32 (0.95)	0.87 (0.84)	–2.04 (1.14)	3.36 (1.48)	2.91 (1.41)

Πηγή: *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, Card and Krueger, 1994

Εστιάζοντας στο τμήμα του πίνακα που βρίσκεται μέσα στο πλαίσιο, βάσει όσων αναφέρθηκαν, τα δύο πρώτα αποτελέσματα της τρίτης στήλης είναι στην ουσία οι συστηματικές διαφορές της πλήρους απασχόλησης, για τις δύο πολιτείες πριν και μετά την εφαρμογή της αύξησης του κατώτατου ωρομισθίου, ενώ τα δύο πρώτα αποτελέσματα της τρίτης γραμμής είναι οι μεταβολές στην πλήρη απασχόληση για κάθε πολιτεία, που οφείλεται στην αλλαγή της πολιτικής. Τέλος το αποτέλεσμα της τρίτης γραμμής και τρίτης στήλης (2.76) είναι η μεταβολή της πλήρους απασχόλησης ανάμεσα στο New Jersey και την Pennsylvania, απαλλαγμένη από συστηματικές μεταβολές και διαχρονικές τάσεις. Είναι στην ουσία ο εκτιμητής των διαφορών σε διαφορές. Εντός των παρενθέσεων στον πίνακα εμφανίζεται το τυπικό σφάλμα του κάθε εκτιμητή.

- **Υπόδειγμα Παλινδρόμησης**

Οι David Card και Alan B. Krueger, θεώρησαν ως ομάδα ελέγχου της έρευνας τα εστιατόρια από την Pennsylvania και ως ομάδα αντιμετώπισης αυτά του New Jersey. Μία απλουστευμένη μορφή του υποδείγματος που χρησιμοποίησαν είναι η παρακάτω:

$$FTE = \beta_0 + \beta_1 NJ + \beta_2 Nov + \beta_3 (NJ \times Nov)$$

όπου FTE η μεταβλητή απόκρισης που περιγράφει το πλήθος των εργαζομένων σε καθεστώς πλήρους απασχόλησης, NJ η ψευδομεταβλητή που λαμβάνει την τιμή ένα αν το εστιατόριο ανήκει στην πολιτεία του New Jersey και μηδέν αν ανήκει στην Pennsylvania, Nov η ψευδομεταβλητή που λαμβάνει την τιμή ένα αν η παρατήρηση προέρχεται από την περίοδο μετά τη μεταβολή και μηδέν από την περίοδο πριν τη

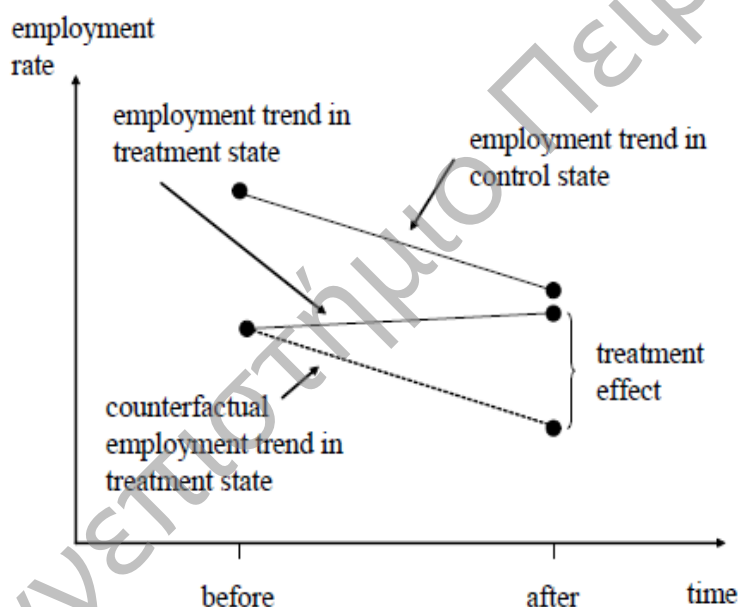
μεταβολή και τέλος $NJ \times Nov$ η αλληλεπίδραση των δύο μεταβλητών. Οι εκτιμήσεις των συντελεστών φαίνονται στον Πίνακα 3.5 και είναι:

- $\hat{\beta}_0 = 23.33$
- $\hat{\beta}_1 = -2.89$
- $\hat{\beta}_2 = -2.16$
- $\hat{\beta}_3 = DD = 2.76$

Επομένως:

$$FTE = 23.33 - 2.89NJ - 2.16Nov + 2.76(NJ \times Nov).$$

Για την αποτελεσματικότητα του υποδείγματος οι David Card και Alan B. Krueger υπέθεσαν ότι οι πορείες της απασχόλησης στις δύο πολιτείες θα ήταν ίδιες (παράλληλες τάσεις) αν δεν υπήρχε παρέμβαση στο κατώτατο ωρομίσθιο όπως φαίνεται και στο Διάγραμμα 3.4.



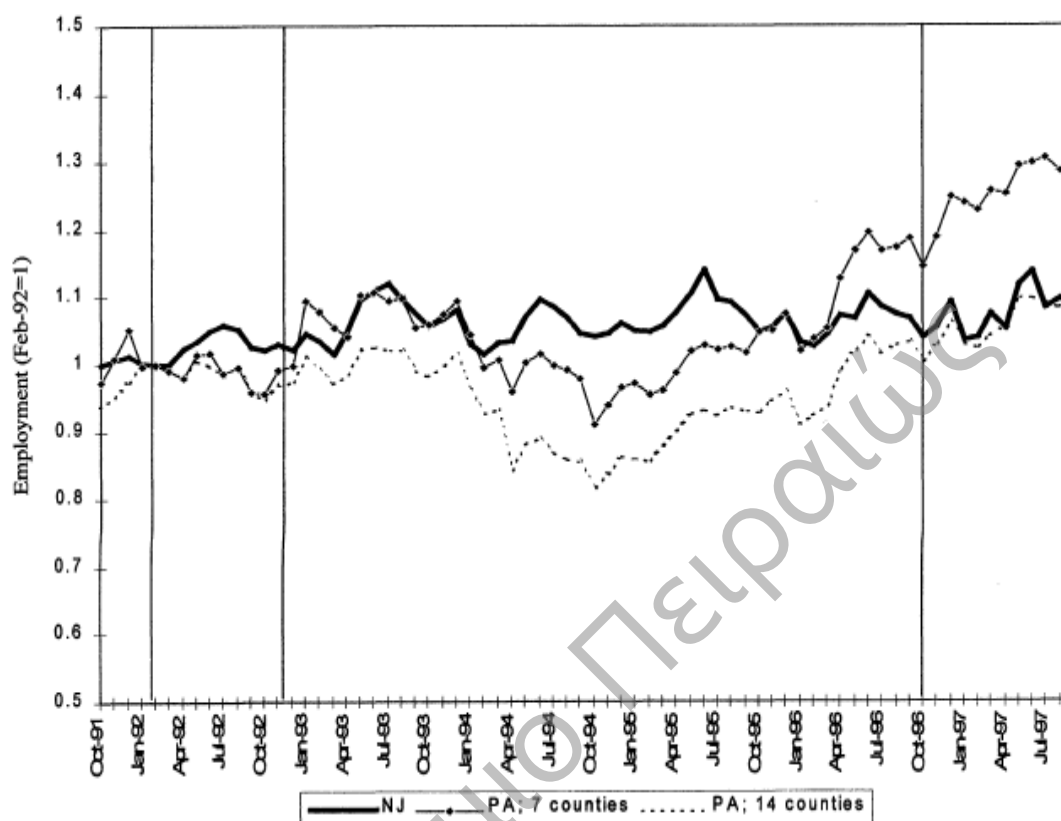
Πηγή: *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, Card and Krueger, 1994

Διάγραμμα 3.4

Υπόθεση ύπαρξης παράλληλης τάσης

Για τη διερεύνηση της υπόθεσης των παράλληλων τάσεων απαιτούνται δεδομένα πολλών περιόδων. Σε μία ενημερωμένη έκδοση της αρχικής τους μελέτης, οι David Card και Alan B. Krueger (2000), συνέλλεξαν δεδομένα μισθοδοσίας πολλών ετών για τις πολιτείες του New Jersey και της Pennsylvania. Τα δεδομένα αυτά εμφανίζονται

στο Διάγραμμα 3.5 και αφορούν τις περιόδους από τον Οκτώβρη του 1991 έως και το Σεπτέμβρη του 1997.



Πηγή: *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, Card and Krueger, 1994

Διάγραμμα 3.5

Μεταβολή απασχόλησης σε New Jersey και Pennsylvania 1991-1997

Οι δύο αριστερές κάθετες γραμμές, προσδιορίζουν την πρώτη και δεύτερη φάση της έρευνάς τους, ενώ η τρίτη την ομοσπονδιακή αύξηση του κατώτατου ωρομισθίου στα \$4.75 τον Οκτώβρη του 1996, γεγονός που επηρέασε την Pennsylvania αλλά όχι το New Jersey.

Όπως και στην αρχική τους μελέτη, από το Φεβρουάριο του 1992 έως και το Νοέμβριο του ίδιου έτους, παρατηρείται μία μικρή μείωση στην απασχόληση για την Pennsylvania και μία μικρή αύξηση για το New Jersey. Ωστόσο, τα δεδομένα αποκαλύπτουν επίσης αρκετά σημαντική διακύμανση της απασχόλησης από έτος σε έτος σε άλλες περιόδους, για τις δύο πολιτείες. Συγκεκριμένα, ενώ τα επίπεδα απασχόλησης στο New Jersey και την Pennsylvania ήταν παρόμοια στο τέλος του 1991, η απασχόληση στην Pennsylvania μειώθηκε σε σχέση με αυτήν του New Jersey

για τα επόμενα τρία έτη, κυρίως πριν από την αύξηση του κατώτατου ωρομίσθιου του 1996. Επομένως, η Pennsylvania μπορεί να μην είναι πολύ καλό μέτρο σύγκρισης της απασχόλησης για το New Jersey ελλείψει μίας μεταβολής στην πολιτική και αντίστροφα.

3.4 Αύξηση Επιδότησης Εργατικού Ατυχήματος και Διάρκεια Άδειας

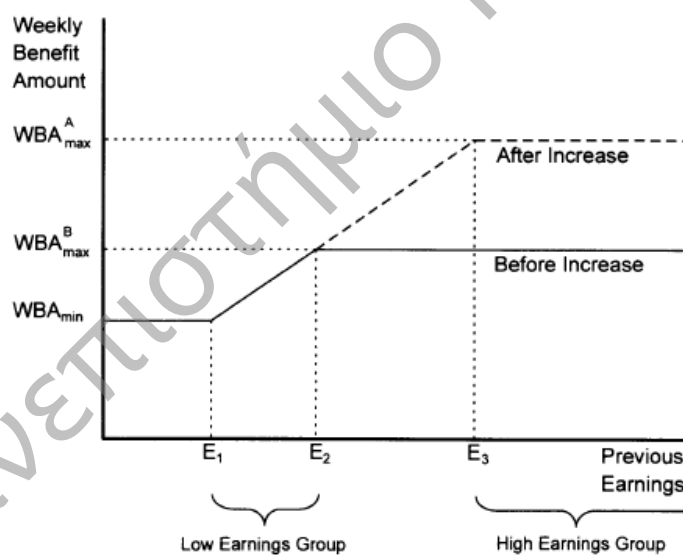
Οι Bruce D.Meyer, W. Kip Viscusi και David L.Durbin, μελέτησαν το 1995 τη χρονική διάρκεια (σε εβδομάδες) που λαμβάνει αποζημίωση ένας εργαζόμενος μετά από εργατικό ατύχημα, για τις πολιτείες του Kentucky και του Michigan. Ο λόγος που έλαβε χώρα αυτή η μελέτη ήταν η θεσμοθέτηση υψηλότερης αποζημίωσης για τους υψηλά αμειβόμενους εργαζόμενους στις δύο αυτές πολιτείες. Θεωρώντας ως ομάδα αντιμετώπισης τους υψηλά αμειβόμενους και ως ομάδα ελέγχου τους χαμηλά αμειβόμενους, συνέλλεξαν για την πολιτεία του Kentucky δείγμα 1,233 υψηλόμισθων και 1,705 χαμηλόμισθων πριν από τη μεταβολή του νόμου, ενώ μετά τη μεταβολή το δείγμα αντιστοιχούσε σε 1,161 υψηλόμισθους και 1,527 χαμηλόμισθους. Αντίστοιχα, για την πολιτεία του Michigan συνέλλεξαν δείγμα 239 υψηλόμισθων και 589 χαμηλόμισθων πριν από τη μεταβολή του νόμου, ενώ μετά τη μεταβολή το δείγμα αντιστοιχούσε σε 219 υψηλόμισθους και 477 χαμηλόμισθους. Η μέθοδος που ακολουθήθηκε είναι αυτή των διαφορών σε διαφορές, αλλά χρησιμοποιήθηκαν δύο υποδείγματα παλινδρόμησης για κάθε πολιτεία. Ένα που η μεταβλητή απόκρισης εμφανιζόταν σε απόλυτα νούμερα και ένα όπου η μεταβλητή απόκρισης εμφανιζόταν σε λογαριθμικές τιμές. Τα αποτελέσματα της μελέτης τους, καθώς και οι αντίστοιχοι πίνακες και τα αντίστοιχα διαγράμματα εμφανίζονται στη συνέχεια.

- **Η μεταβολή του νόμου**

Στις 15 Ιουλίου του 1980, η πολιτεία του Kentucky αύξησε το όριο των εβδομαδιαίων αποδοχών που δικαιούταν ένας εργαζόμενος που βρισκόταν σε αναρρωτική άδεια μετά από εργατικό ατύχημα, από τα \$131 στα \$217, μία μεταβολή ίση με 66%. Σε παρόμοια πρακτική προχώρησε και η πολιτεία του Michigan η οποία την 1^η Ιανουαρίου του 1982 αύξησε τις εβδομαδιαίες αποδοχές για την ίδια περίπτωση,

από τα \$181 στα \$307, αύξηση ίση με 70%. Η μεταβολή αυτή αφορούσε μόνο τους υψηλόμισθους εργαζόμενους και όχι τους χαμηλόμισθους.

Οι Bruce D.Meyer, W. Kip Viscusi και David L.Durbin υπέθεσαν ότι τέτοιες μεταβολές είναι πιθανό να επηρεάσουν τα κίνητρα των εργαζόμενων από τη μείωση της προσοχής τους για αποφυγή τραυματισμών, μέχρι και την επιδίωξη αυτών. Επίσης υψηλότερες παροχές στην αναρρωτική άδεια μπορεί να ωθήσουν τους εργαζόμενους να μείνουν για μεγαλύτερο διάστημα εκτός εργασίας, είτε για να ολοκληρωθεί ορθά η θεραπεία είτε απλά για να έχουν περισσότερο ελεύθερο χρόνο. Από τα παραπάνω λοιπόν δημιουργήθηκε ένα φυσικό πείραμα για τις περιοχές του Kentucky και του Michigan, με ομάδα ελέγχου τους χαμηλόμισθους εργαζόμενους και ομάδα αντιμετώπισης τους υψηλόμισθους εργαζόμενους. Θεωρώντας ότι η αποζημίωση της αναρρωτικής άδειας μεταβάλλεται γραμμικά σε σχέση με το μισθό δημιούργησαν το Διάγραμμα 3.6, όπου με τη συνεχόμενη γραμμή εμφανίζεται το ύψος της εβδομαδιαίας αποζημίωσης πριν την αύξηση, ενώ με τη διακεκομμένη γραμμή το ύψος μετά την αύξηση.



Πηγή: *Workers' Compensation and Injury Duration: Evidence from a Natural Experiment*, Bruce D.Meyer, W. Kip Viscusi and David L.Durbin, 1995

Διάγραμμα 3.6

Ύψος εβδομαδιαίας αποζημίωσης πριν και μετά τη μεταβολή

- Συλλογή δεδομένων και μεταβλητές

Οι μεταβλητές που συνέλεξαν για το σύνολο δεδομένων, ήταν η ημερομηνία τραυματισμού, η διάρκεια του επιδόματος, το συνολικό ιατρικό κόστος, ο

προηγούμενος μισθός, το συνολικό ποσό του εβδομαδιαίου επιδόματος, ο τύπος τραυματισμού, η ηλικία, το φύλο και η οικογενειακή κατάσταση. Στον Πίνακα 3.6, παρουσιάζεται η κατανομή του δείγματος βάσει του είδους τραυματισμού για τις περιοχές του Kentucky και του Michigan. Οι κατηγορίες τραυματισμών που ήταν υψηλές πριν από την αύξηση της αποζημίωσης, παρέμειναν υψηλές για το Kentucky ενώ το Michigan παρουσίασε αξιοσημείωτη μείωση των τραυματισμών των άνω άκρων στους υψηλά αμειβόμενους εργαζόμενους.

Πίνακας 3.6
Κατανομή δείγματος βάσει τραυματισμού

Injury type	Kentucky			Michigan		
	Before increase (1)	After increase (2)	Percentage change (3)	Before increase (4)	After increase (5)	Percentage change (6)
Head, high earnings	4.38 (0.58)	3.36 (0.53)	-23.30 (15.81)	4.18 (1.30)	2.74 (1.10)	-34.52 (33.25)
Head, low earnings	3.40 (0.44)	4.32 (0.52)	27.06 (22.43)	2.72 (0.67)	2.31 (0.69)	-15.11 (32.84)
Neck, high earnings	2.27 (0.42)	3.19 (0.52)	40.34 (34.68)	2.51 (1.01)	1.37 (0.79)	-45.43 (38.25)
Neck, low earnings	0.88 (0.23)	1.31 (0.29)	48.88 (50.58)	1.19 (0.45)	1.26 (0.51)	5.84 (58.52)
Upper extremities, high earnings	23.76 (1.21)	23.51 (1.24)	-1.05 (7.27)	30.96 (2.99)	18.26 (2.61)	-41.01 (10.18)
Upper extremities, low earnings	34.13 (1.15)	33.01 (1.20)	-3.31 (4.80)	33.11 (1.94)	31.45 (2.13)	-5.02 (8.50)
Trunk, high earnings	12.41 (0.94)	12.23 (0.96)	-1.43 (10.75)	14.23 (2.26)	19.63 (2.68)	38.02 (28.92)
Trunk, low earnings	10.26 (0.73)	9.36 (0.75)	-8.76 (9.77)	11.38 (1.31)	12.37 (1.51)	8.74 (18.22)
Low back, high earnings	28.71 (1.29)	28.34 (1.32)	-1.30 (6.39)	21.34 (2.65)	26.03 (2.97)	21.97 (20.56)
Low back, low earnings	24.69 (1.04)	26.20 (1.13)	6.09 (6.40)	25.64 (1.80)	22.85 (1.92)	-10.87 (9.77)
Lower extremities, high earnings	24.98 (1.23)	24.12 (1.26)	-3.45 (6.93)	22.59 (2.71)	27.85 (3.03)	23.28 (19.94)
Lower extremities, low earnings	22.99 (1.02)	21.87 (1.06)	-4.86 (6.24)	19.35 (1.63)	23.69 (1.95)	22.40 (14.39)
Other injuries, high earnings	2.51 (0.45)	4.05 (0.58)	61.02 (36.67)	4.18 (1.30)	2.74 (1.10)	-34.52 (33.25)
Other injuries, low earnings	3.05 (0.42)	3.01 (0.44)	-1.23 (19.69)	3.74 (0.78)	4.61 (0.96)	23.48 (36.45)
Occupational diseases, high earnings	0.97 (0.28)	1.21 (0.32)	23.90 (48.48)	0.00 —	1.37 (0.79)	— —
Occupational diseases, low earnings	0.59 (0.18)	0.92 (0.24)	56.32 (64.49)	2.89 (0.69)	1.47 (0.55)	-49.16 (22.62)
Sample sizes:						
High earnings	1,233	1,161		239	219	
Low earnings	1,705	1,527		589	477	

Notes: Standard errors are in parentheses. The standard errors for the percentage of occupational diseases in Michigan in the high-earnings group cannot be calculated in the usual way.

Πηγή: *Workers' Compensation and Injury Duration: Evidence from a Natural Experiment*, Bruce D. Meyer, W. Kip Viscusi and David L. Durbin, 1995

- **Υπόδειγμα Παλινδρόμησης**

Οι Bruce D.Meyer, W. Kip Viscusi και David L.Durbin για τη μελέτη αυτή θεώρησαν ως ομάδα ελέγχου τους χαμηλόμισθους εργαζόμενους και ως ομάδα αντιμετώπισης τους υψηλόμισθους. Τελικώς οι εξισώσεις παλινδρόμησης είναι της μορφής:

$$durat = \beta_0 + \beta_1 wage + \beta_2 after + \beta_3(wage \times after)$$

και

$$\log(durat) = \beta_0 + \beta_1 wage + \beta_2 after + \beta_3(wage \times after)$$

όπου η μεταβλητή απόκρισης *durat* είναι η διάρκεια της άδειας του κάθε εργαζόμενου λόγω τραυματισμού σε εβδομάδες, ενώ $\log(durat)$ ο λογάριθμος αυτής, η ψευδομεταβλητή *wage* λαμβάνει την τιμή ένα αν ο εργαζόμενος ανήκει στην ομάδα των υψηλόμισθων και την τιμή μηδέν αν ανήκει στην ομάδα των χαμηλόμισθων, η ψευδομεταβλητή *after* λαμβάνει την τιμή ένα αν η μεταβλητή απόκρισης αναφέρεται στην περίοδο μετά την μεταβολή της νομοθεσίας και την τιμή μηδέν αν αναφέρεται πριν από αυτήν και τέλος η μεταβλητή *wage × after*, αναφέρεται στην αλληλεπίδραση των δύο αυτών μεταβλητών. Στον Πίνακα 3.7 παρουσιάζονται τα αποτελέσματα της εξίσωσης παλινδρόμησης για τις πολιτείες των Kentucky και Michigan. Στις στήλες (1) και (2) παρουσιάζεται η μέση διάρκεια της άδειας για τους υψηλόμισθους εργαζόμενους πριν και μετά την αύξηση και στη στήλη (5) η μεταβολή που επήλθε λόγω της νέας πολιτικής. Στις στήλες (3) και (4) παρουσιάζεται η μέση διάρκεια της άδειας για τους χαμηλόμισθους πριν και μετά τη αλλαγή της πολιτικής, ενώ η στήλη (6) δείχνει τη μεταβολή της μέσης διάρκειας της άδειας τους. Η τελευταία στήλη του πίνακα (7) είναι ο συντελεστή των διαφορών σε διαφορές, απαλλαγμένος από διαχρονικές τάσεις και συστηματικές μεταβολές ενώ στις παρενθέσεις του πίνακα παρουσιάζονται τα τυπικά σφάλματα των παραμέτρων.

Οι γραμμές που βρίσκονται εντός του πρώτου πλαισίου, παρουσιάζουν τα αποτελέσματα για τη μέση διάρκεια της άδειας για το Kentucky και το Michigan, ενώ το δεύτερο πλαίσιο αναφέρεται στα αποτελέσματα του λογάριθμου της μέσης διάρκειας της άδειας. Στη στήλη (7) οι συντελεστές των διαφορών σε διαφορές είναι όλοι θετικοί, φανερώνοντας έτσι θετική επίδραση της πολιτικής στη διάρκεια της άδειας. Υπολογίζοντας τις τιμές των στατιστικών συναρτήσεων των συντελεστών των διαφορών σε διαφορές προκύπτει πως μόνο η λογαριθμική διάρκεια της άδειας είναι στατιστικά σημαντική και αυτό μόνο για την περίπτωση του Kentucky.

Πίνακας 3.7

Εκτίμηση υποδείγματος διαφορών σε διαφορές

Variable	High earnings		Low earnings		Differences		Difference in differences
	Before increase (1)	After increase (2)	Before increase (3)	After increase (4)	[(2)-(1)] (5)	[(4)-(3)] (6)	[(5)-(6)] (7)
Mean duration (weeks)							
Kentucky	11.16 (0.83)	12.89 (0.83)	6.25 (0.30)	7.01 (0.41)	1.72 (1.17)	0.76 (0.51)	0.96 (1.28)
Michigan	14.76 (2.25)	19.42 (2.67)	10.94 (1.09)	13.64 (1.56)	4.66 (3.49)	2.70 (1.90)	1.96 (3.97)
Median duration (weeks)							
Kentucky	4.00 (0.14)	5.00 (0.20)	3.00 (0.11)	3.00 (0.12)	1.00 (0.25)	0.00 (0.16)	1.00 (0.29)
Michigan	5.00 (0.45)	7.00 (0.67)	4.00 (0.22)	4.00 (0.28)	2.00 (0.81)	0.00 (0.35)	2.00 (0.89)
75th percentile, duration (weeks)							
Kentucky	8.00 (0.28)	10.00 (0.45)	7.00 (0.21)	7.00 (0.24)	2.00 (0.53)	0.00 (0.32)	2.00 (0.62)
Michigan	10.00 (0.74)	14.00 (1.88)	8.50 (0.54)	9.00 (0.57)	4.00 (2.03)	0.50 (0.79)	3.50 (2.17)
Mean of log duration							
Kentucky	1.38 (0.04)	1.58 (0.04)	1.13 (0.03)	1.13 (0.03)	0.20 (0.05)	0.01 (0.04)	0.19 (0.07)
Michigan	1.58 (0.09)	1.87 (0.10)	1.41 (0.06)	1.51 (0.06)	0.29 (0.13)	0.10 (0.08)	0.19 (0.16)
Mean medical cost (dollars)							
Kentucky	3,298.96 (1,885.93)	1,688.67 (116.59)	878.20 (78.07)	1,155.98 (157.26)	-1,610.29 (1,889.53)	277.78 (175.57)	-1,888.07 (1,897.67)
Michigan	2,229.41 (293.93)	2,585.23 (364.76)	1,538.22 (188.49)	2,017.65 (304.83)	355.82 (468.45)	479.43 (358.40)	-123.61 (589.83)
Median medical cost (dollars)							
Kentucky	393.51 (19.29)	411.49 (22.72)	238.96 (8.48)	254.40 (9.11)	17.98 (29.80)	15.44 (12.44)	2.55 (32.30)
Michigan	689.73 (77.30)	765.00 (134.53)	390.63 (32.80)	435.00 (33.09)	75.27 (155.16)	44.38 (46.59)	30.89 (162.00)
75th percentile, cost (dollars)							
Kentucky	1,335.71 (103.08)	1,686.40 (122.95)	864.94 (72.24)	867.53 (69.78)	350.69 (160.45)	2.59 (100.44)	348.10 (189.29)
Michigan	2,284.60 (178.51)	2,379.00 (284.80)	1,383.93 (155.69)	1,822.00 (145.49)	94.40 (336.12)	438.07 (213.09)	-343.67 (397.98)
Mean of log cost							
Kentucky	6.09 (0.05)	6.24 (0.05)	5.61 (0.04)	5.69 (0.04)	0.15 (0.07)	0.08 (0.06)	0.07 (0.09)
Michigan	6.56 (0.11)	6.59 (0.13)	5.85 (0.09)	6.10 (0.09)	0.03 (0.17)	0.25 (0.12)	-0.22 (0.21)

Πηγή: *Workers' Compensation and Injury Duration: Evidence from a Natural Experiment*, Bruce D.Meyer, W. Kip Viscusi and David L.Durbin, 1995

Κοινή συνισταμένη στις μελέτες των David Card και Alan B.Krueger, καθώς και σε αυτή των Bruce D.Meyer, W. Kip Viscusi και David L.Durbin αποτελεί το γεγονός ότι η μέθοδος των διαφορών σε διαφορές χρησιμοποιήθηκε για να εξαχθούν συμπεράσματα αντίθετα από τις εκτιμήσεις και τις θέσεις των ερευνητών της εποχής. Το γεγονός ότι μέσω αυτής, δίνεται η δυνατότητα στον εκάστοτε ερευνητή να απαλλαγεί από τις συστηματικές διαφορές των ομάδων ελέγχου και αντιμετώπισης, καθώς και από διαχρονικές τάσεις, την καθιστά πολύ σημαντική και χρήσιμη στις εμπειρικές μελέτες.

3.5 Ανακεφαλαίωση

Στα πλαίσια του παρόντος κεφαλαίου έγινε προσπάθεια να γίνει πιο κατανοητή στον αναγνώστη η χρήση της μεθόδου των διαφορών σε διαφορές, μέσω της παρουσίασης εμπειρικών μελετών όπου έχει γίνει χρήση της μεθόδου. Αρχικά παρουσιάστηκε η μελέτη των Marie Louise Obenauer και Bertha Marie von der Nienburg που πραγματοποιήθηκε το 1915 και στην οποία ερευνήθηκαν οι επιπτώσεις της θεσμοθέτησης ενός κατώτατου εβδομαδιαίου μισθού στη γυναικεία απασχόληση στο Oregon. Στη συνέχεια έγινε αναφορά σε μία από τις πιο διάσημες μελέτες που έγινε χρήση της μεθόδου των διαφορών σε διαφορές αυτή των David Card και Alan B.Krueger (1994), οι οποίοι μελετώντας τις επιπτώσεις της αύξησης του ωρομίσθιου, από τα \$4.25 στα \$5.05 στην πολιτεία του New Jersey, κατέληξαν στο συμπέρασμα ότι η αύξηση αυτή οδήγησε σε αύξηση της απασχόλησης στην περίπτωση των ταχυφαγείων, σε αντίθεση με τις εκτιμήσεις των ερευνητών της εποχής οι οποίοι πίστευαν ότι μία ενδεχόμενη αύξηση στους μισθούς θα οδηγούσε σε απώλεια θέσεων εργασίας. Τέλος, έγινε αναφορά στη μελέτη των Bruce D.Meyer, W.Kip Viscusi και David L.Durbin του 1995, οι οποίοι μελέτησαν τις επιπτώσεις της αύξησης της αποζημίωσης στη διάρκεια της αναρρωτικής άδειας, μετά από εργατικό ατύχημα.

ΚΕΦΑΛΑΙΟ 4
ΟΙΚΟΝΟΜΙΚΗ ΚΡΙΣΗ ΚΑΙ ΑΓΟΡΑ
ΑΥΤΟΚΙΝΗΤΟΥ
ΕΜΠΕΙΡΙΚΗ ΔΙΕΡΕΥΝΗΣΗ

4.1 Εισαγωγή

Το έτος 2007 είναι συνυφασμένο με την απαρχή μίας πρωτόγνωρης οικονομικής κρίσης, η οποία ξεκίνησε στις Η.Π.Α και σε σύντομο χρονικό διάστημα, μετατράπηκε σε κρίση του παγκόσμιου χρηματοπιστωτικού συστήματος και μεταφέρθηκε στην πραγματική οικονομία με ταχύ ρυθμό. Στην Ευρωπαϊκή Ήπειρο, χώρες όπως η Ιρλανδία, η Ελλάδα, η Ισπανία, η Ιταλία και η Πορτογαλία που μέχρι πρότινος εμφάνιζαν θετικούς ρυθμούς ανάπτυξης, εμφάνισαν μεγάλα ποσοστά ελλειμάτων σε σχέση με το ΑΕΠ τους και με στόχο τη μη χρεοκοπία τους, δημιουργήθηκε η ανάγκη πρόσθετης χρηματοδότησης και επιτήρησης για αυτές. Την ίδια στιγμή, τα ποσοστά ανεργίας αυξήθηκαν και η ζήτηση των καταναλωτικών προϊόντων μειώθηκε. Μέσα σε αυτό το ευμετάβλητο οικονομικό περιβάλλον, μεγάλο μέρος του καταναλωτικού κοινού αναθεώρησε τις προτιμήσεις του και άρχισε να λειτουργεί ορθολογικά, επιλέγοντας οικονομικότερα αγαθά. Το φαινόμενο αυτό παρατηρήθηκε σε όλους τους κλάδους της οικονομικής δραστηριότητας, καθώς και σε αυτόν του αυτοκινήτου, ο οποίος κλονίστηκε σύν τοις άλλοις και από την υπερφορολόγηση των καυσίμων.

Σκοπός αυτού του κεφαλαίου είναι η διερεύνηση των επιπτώσεων της ευρωπαϊκής οικονομικής κρίσης και της συνεχώς αυξανόμενης τιμής του πετρελαίου, στις καταναλωτικές προτιμήσεις, ως προς την αγορά καινούργιου αυτοκινήτου με βενζινοκινητήρα (*Gasoline*) ή τον πιο οικονομικό σε κατανάλωση πετρελαιοκινητήρα (*Diesel*), για ένα σύνολο χωρών της Ευρώπης. Πιο συγκεκριμένα, η μελέτη αναφέρεται σε δείγμα μορφής πάνελ και προέρχεται από τις χώρες Αυστρία, Βέλγιο, Δανία, Φινλανδία, Γαλλία, Γερμανία, Ελλάδα, Ιρλανδία, Ιταλία, Λουξεμβούργο, Ολλανδία, Πορτογαλία, Ισπανία, Σουηδία, Ηνωμένο Βασίλειο που ανήκουν στην Ευρωπαϊκή Ένωση των 15 χωρών (*EU15*) και τις Ισλανδία, Νορβηγία και Ελβετία που ανήκουν στην Ευρωπαϊκή Ζώνη Ελεύθερων Συναλλαγών (*EFTA3*). Για την εξαγωγή των

εμπειρικών αποτελεσμάτων χρησιμοποιείται το στατιστικό πακέτο SPSS, ενώ η οπτικοποίηση των δεδομένων πραγματοποιείται με χρήση του λογισμικού Excel.

Αρχικά, παρουσιάζεται η ιστορική αναδρομή της ευρωπαϊκής οικονομικής κρίσης, ενώ στη συνέχεια παρουσιάζονται τα δεδομένα που χρησιμοποιούνται στην εμπειρική ανάλυση, μέσα από πίνακες και διαγράμματα. Κατόπιν λαμβάνει χώρα η ανάλυση των δεδομένων μέσα από την παραλλαγή της μεθόδου των διαφορών σε διαφορές για δεδομένα τύπου πάνελ, ενώ, τέλος ελέγχεται αν οι πωλήσεις των δειγμάτων κατανομονται με όμοιο τρόπο μέσα από στατιστικούς ελέγχους, έτσι ώστε να εξασφαλιστεί η συνθήκη της παράλληλης τάσης.

4.2 Ιστορική Αναδρομή της Οικονομικής Κρίσης

Η διεθνής οικονομία βρέθηκε αντιμέτωπη με τη μεγαλύτερη ύφεση μετά το τέλος του Β' Παγκοσμίου Πολέμου. Η χρηματοοικονομική κρίση, που ξεκίνησε στην αγορά στεγαστικών δανείων χαμηλής εξασφάλισης (*subprime mortgages*) των ΗΠΑ το καλοκαίρι του 2007 μετατράπηκε ραγδαία σε κρίση του παγκόσμιου χρηματοπιστωτικού συστήματος και μεταφέρθηκε στην πραγματική οικονομία με ταχύ ρυθμό.

Μια σειρά από γεγονότα, οδήγησαν τους παγκόσμιους ρυθμούς ανάπτυξης σε απότομη πτώση και εκτόξευσαν την ανεργία και την ανασφάλεια στο σύνολο σχεδόν των χωρών. Καθώς η κρίση εκτυλισσόταν, η ένταση και η έκτασή της αύξαναν συνεχώς, αναγκάζοντας κυβερνήσεις, κεντρικές τράπεζες, αναλυτές, επενδυτές, επιχειρηματίες και καταναλωτές να αναθεωρούν συνεχώς τις αντιλήψεις τους και τις προσδοκίες τους. Στην προσπάθεια άμεσης αντιμετώπισης της κρίσης ελήφθησαν μια σειρά από μέτρα και πολιτικές, όπως η κρατικοποίηση τραπεζών. Οι προηγούμενες αντιλήψεις για μια νέα εποχή χαμηλής μεταβλητότητας των οικονομικών μεγεθών, συνεχούς ευημερίας και αυτορρύθμισης των αγορών, κατέρρευσαν μπροστά στη ραγδαία επιδείνωση της κατάστασης. Μόνο μία φορά στο παρελθόν η παγκόσμια οικονομία βρέθηκε μπροστά σε τέτοιο κίνδυνο, τη δεκαετία του 1930.

Η πρώτη φούσκα έσπασε στην αγορά κατοικιών στις ΗΠΑ το καλοκαίρι του 2007. Σε μεγάλο βαθμό, οι κατοικίες αυτές είχαν αγοραστεί με ενυπόθηκα δάνεια μεγάλης μόχλευσης και χαμηλής πιστοληπτικής διαβάθμισης. Στη συνέχεια, ακολούθησαν οι αγορές όπου διαπραγματεύονταν προϊόντα τα οποία προέρχονταν από τιτλοποιήσεις

στεγαστικών δανείων. Αργότερα ακολούθησαν φούσκες σε αγορές πιο σύνθετων προϊόντων, όπως τα εξασφαλισμένα ομόλογα χρέους (*CDOs – Collateralized Debt Obligations*) και οι συμβάσεις ανταλλαγής κινδύνου αθέτησης (*CDSs – Credit Default Swaps*).

Η ικανότητα των τραπεζών να απορροφήσουν κινδύνους βρέθηκε σε ιστορικά χαμηλά επίπεδα. Λόγω ανεπαρκούς κεφαλαιακής βάσης, αλλά και λόγω του τεράστιου βραχυπρόθεσμου δανεισμού, οι «φούσκες» στις αγορές φυσικού και χρηματοοικονομικού κεφαλαίου συνέχισαν να σπάνε με σημείο κορύφωσης την πτώχευση της Lehman Brothers και την κατάρρευση μεγάλων τραπεζικών και ασφαλιστικών οργανισμών το Φθινόπωρο του 2008. Ο άνευ προηγουμένου κλονισμός του διεθνούς χρηματοοικονομικού συστήματος, οδήγησε σε κρίση εμπιστοσύνης, κυρίως μεταξύ πιστωτικών οργανισμών που ως κύρια πηγή χρηματοδότησης, είχαν τον βραχυπρόθεσμο δανεισμό, η οποία στη συνέχεια εκδηλώθηκε ως κρίση ρευστότητας.

Ιδίως κατά το τελευταίο τρίμηνο του 2008 παρατηρήθηκε κατάκόρυφη αύξηση του κόστους χρήματος για τις παρατηρήσεις της πραγματικής οικονομίας, ενώ σε ορισμένες περιπτώσεις το περιθώριο του επιτοκίου δανεισμού των πιστωτικών ιδρυμάτων πάνω από το βραχυπρόθεσμο επιτόκιο δανεισμού του αντίστοιχου δημοσίου επταπλασιάστηκε. Η χρηματοοικονομική κρίση πέρασε στην πραγματική οικονομία, μέσω της απροθυμίας των πιστωτικών ιδρυμάτων να χορηγήσουν δάνεια και μέσω της καταστροφής του πλούτου νοικοκυριών και επιχειρήσεων. Το εξωτερικό εμπόριο κατέρρευσε και για πρώτη φορά στα τελευταία τριάντα χρόνια, μειώθηκε η κατανάλωση. Συνεπεία τούτων ήταν η εντυπωσιακή μείωση του ΑΕΠ παγκοσμίως, η μείωση της απασχόλησης και η αύξηση της ανεργίας, όπως παρουσιάζεται και στους Πίνακα 4.1 και Πίνακα 4.2.

Πίνακας 4.1

Ρυθμός ανάπτυξης

	2007	2008	2009	2010
Παγκόσμια Οικονομία	5,20%	3,00%	-1,10%	3,10%
Ανεπτυγμένες Οικονομίες	2,70%	0,60%	-3,40%	1,30%
ΗΠΑ	2,10%	0,40%	-2,70%	1,50%
Ευρωζώνη	2,70%	0,70%	-4,20%	0,30%
ΕΕ	3,10%	1,00%	-4,20%	0,50%

Πηγή: Διεθνές Νομισματικό Ταμείο, *World Economic Outlook, Oct '09*

Πίνακας 4.2

Ποσοστό ανεργίας % του εργατικού δυναμικού

	2007	2008	2009	2010
ΗΠΑ	4,60%	5,80%	9,30%	10,10%
Ευρωζώνη	7,40%	7,50%	10,00%	12,00%
ΕΕ	7,50%	7,50%	9,50%	10,70%

Πηγή: *OECD Economic Outlook, Nov. '09*

Όταν το Κραχ του 2008 χτύπησε τις ΗΠΑ, η Ευρωπαϊκή οικονομία δέχτηκε ισχυρό πλήγμα. Με την κατάρρευση των καταναλωτικών αγορών των Ηνωμένων Πολιτειών και του μηχανισμού ανακύκλωσης πλεονασμάτων, η Ευρώπη όχι μόνο έχασε μια σημαντική πηγή ζήτησης για τα προϊόντα της, αλλά συνειδητοποίησε και ότι οι ίδιες οι τράπεζές της αντιμετώπιζαν την καταστροφή, καθώς ήταν εκτεθειμένες στα CDO των ΗΠΑ. Παρά τους ισχυρισμούς των Ευρωπαίων ότι η κρίση δεν θα άγγιζε την Ευρώπη και ότι οι ευρωπαϊκές τράπεζες δεν ήταν εκτεθειμένες σε κίνδυνο, έγινε σύντομα φανερό ότι ίσχυε ακριβώς το αντίθετο. Μέσα σε μία χρονιά (2008-2009), το ΑΕΠ των ευρωπαϊκών κρατών μειώνεται και οι οικονομίες τους βρίσκονται σε ύφεση. Την ίδια στιγμή τα πενταετή ασφάλιστρα κινδύνου CDS από τα τέλη του 2009 έως τις αρχές του 2011 για τις χώρες Ιταλία, Ιρλανδία, Ελλάδα, Πορτογαλία και Ισπανία, αυξάνονται δραματικά καθιστώντας τις χώρες αυτές αδύναμους κρίκους για τη ζώνη του Ευρώ. Ο κίνδυνος διάλυσης της ευρωζώνης είναι υπαρκτός και δημιουργείται η ανάγκη διαρθρωτικών αλλαγών, πρόσθετης χρηματοδότησης και επιτήρησης για τις χώρες αυτές. Τα παραπάνω συνοψίζονται στον Πίνακα 4.3 και στο Διάγραμμα 4.1.

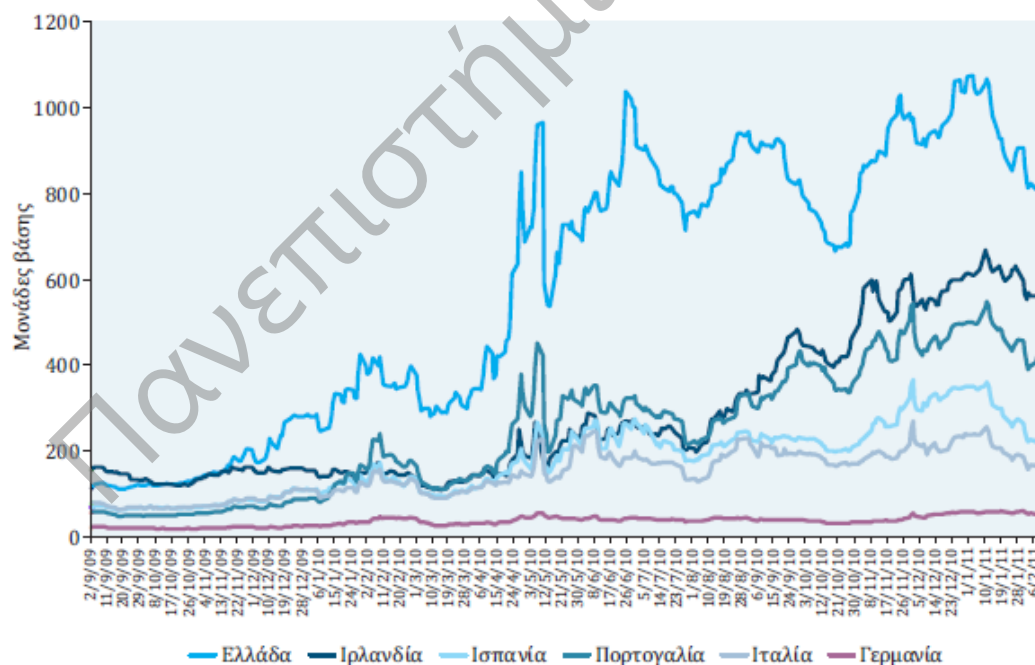
Μέσα σε αυτό το ευμετάβλητο και αβέβαιο οικονομικό περιβάλλον, οι κυβερνήσεις με στόχο την αύξηση των δημοσίων εσόδων, υπερφορολογούν νοικοκυριά και επιχειρήσεις, τα οποία την ίδια στιγμή έρχονται αντιμέτωπα με την αύξηση της ανεργίας και την άρνηση των τραπεζών για χορήγηση δανείων. Το τέλος του 2011 βρίσκει την Ευρώπη να επιστρέφει στο 2008, με μία νέα κρίση τραπεζών η οποία απειλεί το ευρώ με τελική κατάρρευση.

Πίνακας 4.3

Ρυθμός ανάπτυξης Ευρωπαϊκών χωρών

	2008	2009	2010	2011
Αυστρία	1,40%	-3,80%	1,80%	2,80%
Βέλγιο	1,00%	-2,80%	2,30%	1,80%
Δανία	-0,80%	-5,70%	1,40%	1,10%
Φινλανδία	0,30%	-8,50%	3,40%	2,80%
Γαλλία	-0,10%	-3,10%	1,70%	2,00%
Γερμανία	1,10%	-5,10%	4,00%	3,30%
Ελλάδα	-0,20%	-3,10%	-4,90%	-7,10%
Ιρλανδία	-2,20%	-6,40%	-1,10%	2,20%
Ιταλία	-1,20%	-5,50%	1,70%	0,40%
Λουξεμβούργο	-0,70%	-5,60%	3,10%	1,90%
Ολλανδία	1,80%	-3,70%	1,50%	0,90%
Πορτογαλία	0,00%	-2,90%	1,90%	-1,30%
Ισπανία	0,90%	-3,80%	-0,20%	0,10%
Σουηδία	-0,60%	-5,00%	6,60%	2,90%
Ηνωμένο Βασίλειο	-0,80%	-5,20%	1,70%	1,10%
Ισλανδία	1,20%	-6,60%	-4,10%	2,70%
Νορβηγία	0,10%	-1,60%	0,50%	1,30%
Ελβετία	2,20%	-1,90%	3,00%	1,80%

Πηγή: Eurostat, Real GDP Growth Rate, 2008-2011

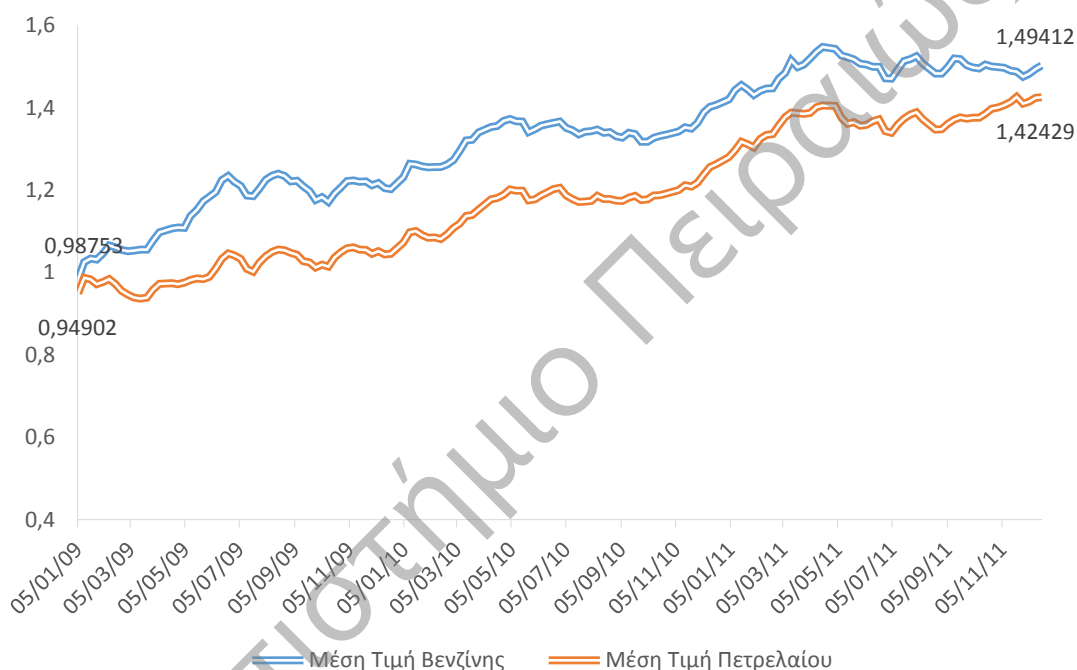


Πηγή: Bloomberg

Διάγραμμα 4.1

Πενταετή CDS, Σεπτέμβριος '09 – Φεβρουάριος '11

Δεδομένης της χρηματοπιστωτικής κρίσης και των επιπτώσεών της, οι καταναλωτές αναθεωρούν τις προτιμήσεις τους και λειτουργούν ορθολογικά, προτιμώντας προϊόντα που τους παρέχουν μεγαλύτερη οικονομία, βραχυπρόθεσμα ή μακροπρόθεσμα. Δε θα μπορούσε να απουσιάζει από τη μεταβολή των καταναλωτικών προτιμήσεων ο κλάδος του αυτοκινήτου, ο οποίος κλονίζεται συν τοις άλλοις και από την υπερφορολόγηση των καυσίμων και κατ' επέκταση την αύξηση των τιμών τους. Στο Διάγραμμα 4.2 παρουσιάζεται η μέση τιμή πώλησης σε ευρώ ανά λίτρο της βενζίνης και του πετρελαίου κίνησης, για την Ευρωπαϊκή Ένωση, όπως καταγράφηκε από την Eurostat για τα έτη 2009 έως και 2011.



Πηγή: Eurostat, Oil Bulletin Prices History, 2009 - 2011

Διάγραμμα 4.2

Μέση τιμή καυσίμων στην Ευρωπαϊκή ήπειρο σε ευρώ ανά λίτρο

Όπως γίνεται αντιληπτό από το Διάγραμμα 4.2 η μέση τιμή της βενζίνης στις αρχές του 2009 ήταν ίση με €0,987 ανά λίτρο και η αντίστοιχη του πετρελαίου ήταν ίση με €0,95 ανά λίτρο, ενώ το τέλος του 2011 βρίσκει την τιμή της βενζίνης στα €1,494 ανά λίτρο και του πετρελαίου στα €1,424 ανά λίτρο. Με απλούς υπολογισμούς, ένα βενζινοκίνητο όχημα μεσαίας κατηγορίας με χωρητικότητα δεξαμενής καυσίμου 45 λίτρων, χρειαζόταν το 2009 ποσό ίσο με €44,42 για να είναι πλήρες, ενώ στο τέλος του 2011 το αντίστοιχο ποσό έφτανε τα €67,23. Αντίστοιχα, το 2009, για ένα πετρελαιοκίνητο

όχημα ίδιων χαρακτηριστικών με αυτό της βενζίνης, απαιτούνταν €42,75 για να είναι πλήρες, ενώ στο τέλος του 2011 το ποσό αυτό, έφτανε τα €64,08.

4.3 Παρουσίαση Δεδομένων

Στην παρούσα μελέτη, διερευνάται το αν οι καταναλωτές από τις χώρες Αυστρία, Βέλγιο, Δανία, Φινλανδία, Γαλλία, Γερμανία, Ελλάδα, Ιρλανδία, Ιταλία, Λουξεμβούργο, Ολλανδία, Πορτογαλία, Ισπανία, Σουηδία, Ηνωμένο Βασίλειο (*EU15*) και Ισλανδία, Νορβηγία και Ελβετία (*EFTA3*) επέλεξαν το 2011 το καινούργιο τους αυτοκίνητο να είναι εφοδιασμένο με βενζινοκινητήρα (*Gasoline*) ή τον πιο οικονομικό σε κατανάλωση αλλά και πιο ακριβό σε αξία κτίσης πετρελαιοκινητήρα (*Diesel*), εξαιτίας της χρηματοπιστωτικής κρίσης. Βάσει του Ευρωπαϊκού κανονισμού υπ' αριθμόν 715/2007 του Ευρωπαϊκού Κοινοβουλίου και του Συμβουλίου της 20ής Ιουνίου 2007, οι αυτοκινητοβιομηχανίες που επιθυμούν να διαθέσουν τα οχήματά τους στην Ευρωπαϊκή Ένωση, είναι υποχρεωμένες να διαθέτουν μεγάλα ποσά στην έρευνα και ανάπτυξη κινητήρων πετρελαίου και βενζίνης, που εκπέμπουν μικρή ποσότητα ρύπων στην ατμόσφαιρα. Η έγκριση για την καταλληλότητα του κινητήρα είναι γνωστή και ως Euro 5 και Euro 6. Ιδιαίτερα οι πετρελαιοκινητήρες, λόγω των αυστηρών προδιαγραφών που έχουν θεσπιστεί, έχουν τέτοια τεχνολογική εξάρση τα τελευταία χρόνια, που τους έχει ανακηρύξει ως τους πιο οικονομικούς και φιλικούς προς το περιβάλλον και τον άνθρωπο, κινητήρες εσωτερικής καύσης. Παρόλα αυτά η απόκτηση ενός πετρελαιοκίνητου οχήματος, είναι κατά κάποιες χιλιάδες ευρώ πιο ακριβή από το αντίστοιχο βενζινοκίνητο, με αποτέλεσμα σε μεσοπρόθεσμο ορίζοντα η απόκτηση του να μην είναι προσιτή.

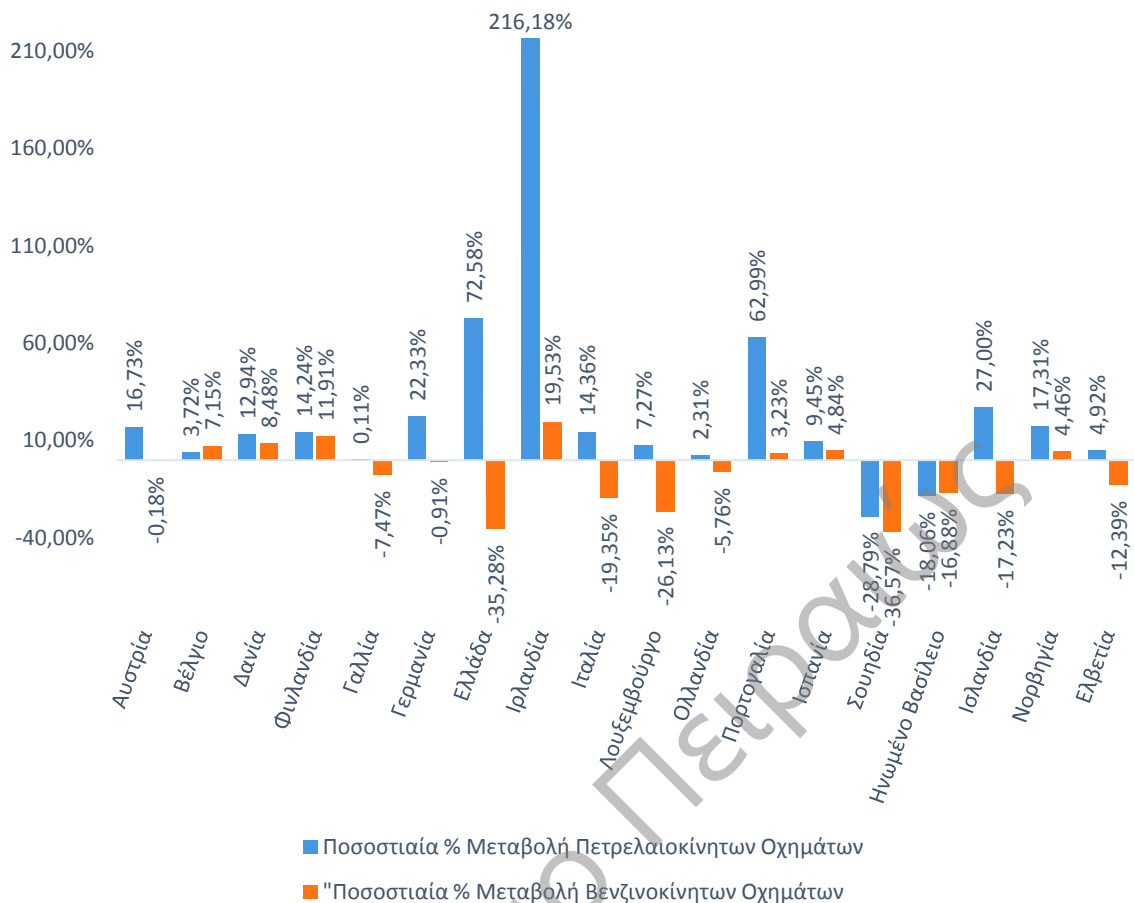
Η εμπειρική ανάλυση πραγματοποιήθηκε χρησιμοποιώντας δεδομένα μορφής πάνελ (*panel data*) για τις παραπάνω χώρες, για τα έτη 2010 και 2011, με ομάδα ελέγχου τα οχήματα που είναι εφοδιασμένα με βενζινοκινητήρα, ενώ ομάδα αντιμετώπισης τα οχήματα που είναι εφοδιασμένα με πετρελαιοκινητήρα, όπως τα κατέγραψε η Ένωση Ευρωπαίων Κατασκευαστών Αυτοκινήτων (*ACEA – European Automobile Manufacturers Association*). Πιο συγκεκριμένα, τα δεδομένα αυτά αφορούν τις ετήσιες πωλήσεις αυτοκινήτων, στις δύο κατηγορίες καυσίμου και παρουσιάζονται στον Πίνακα 4.4, ενώ παράλληλα στο Διάγραμμα 4.3 παρουσιάζεται η ποσοστιαία μεταβολή των πωλήσεων στις δύο κατηγορίες οχημάτων για τα έτη 2010 και 2011.

Πίνακας 4.4
Ετήσιες πωλήσεις αυτοκινήτων

	2010		2011	
	Πετρέλαιο	Βενζίνη	Πετρέλαιο	Βενζίνη
Αυστρία	166.581	161.982	194.455	161.690
Βέλγιο	415.431	131.909	430.875	141.336
Δανία	70.189	83.398	79.270	90.474
Φινλανδία	44.549	62.797	50.892	70.279
Γαλλία	1.594.182	657.487	1.595.862	608.367
Γερμανία	1.221.913	1.694.346	1.494.782	1.678.852
Ελλάδα	5.660	135.841	9.768	87.912
Ιρλανδία	674	2.432	2.131	2.907
Ιταλία	55.013	33.433	62.915	26.963
Λουξεμβούργο	900.365	1.061.214	965.856	783.883
Ολλανδία	37.394	12.332	38.259	11.622
Πορτογαλία	96.513	386.054	157.304	398.540
Ισπανία	95.688	32.066	104.727	33.618
Σουηδία	149.944	73.520	106.769	46.635
Ηνωμένο Βασίλειο	693.303	288.712	568.060	239.991
Ισλανδία	147.449	142.235	187.260	117.724
Νορβηγία	88.321	204.132	103.609	213.237
Ελβετία	936.220	1.094.626	982.274	958.979

Πηγή: ACEA – European Automobile Manufacturers Association, 2010 – 2011

Από το Διάγραμμα 4.3, γίνεται αντιληπτό το γεγονός, ότι οι πωλήσεις των βενζινοκίνητων οχημάτων μεταβλήθηκαν από ελάχιστα έως και καθόλου σε ορισμένες χώρες, ενώ των πετρελαιοκίνητων σημείωσαν ιδιαίτερη μεταβολή. Ιδιαίτερο ενδιαφέρον παρουσιάζουν οι χώρες Ιρλανδία, Ελλάδα και Πορτογαλία, οι οποίες κατά τη διάρκεια της οικονομικής κρίσης, αύξησαν κατά μεγάλο ποσοστό το στόλο των πετρελαιοκίνητων οχημάτων, ενώ στην αντίπερα όχθη μόνο το Βέλγιο εμφάνισε μεγαλύτερη μεταβολή υπέρ των βενζινοκίνητων οχημάτων. Το διάγραμμα αυτό δίνει μία πρώτη και κατατοπιστική εικόνα του πως μεταβλήθηκαν οι καταναλωτικές προτιμήσεις στον κλάδο της αυτοκινητοβιομηχανίας στην Ευρωπαϊκή ήπειρο τα έτη 2010 και 2011.



Πηγή: ACEA – European Automobile Manufacturers Association, 2010 - 2011

Διάγραμμα 4.3

Ποσοστιαία % μεταβολή πωλήσεων ανά κατηγορία οχήματος

4.4 Εμπειρική Ανάλυση Δεδομένων

Σκοπός στην παρούσα μελέτη είναι όχι μόνο η εξαγωγή στατιστικών συμπερασμάτων μέσω της μεθόδου των διαφορών σε διαφορές με χρήση δεδομένων πάνελ, αλλά και η ανάδειξη της σημαντικότητας της συγκεκριμένης μεθόδου, έναντι άλλων που χρησιμοποιούν απλώς διαστρωματικά δεδομένα. Πριν από την παρουσίαση των αποτελεσμάτων της ανάλυσης μέσω της μεθόδου, θα ήταν χρήσιμο να παρουσιαστούν οι στατιστικές αναλύσεις για το πώς κινήθηκαν οι καταναλωτικές προτιμήσεις στον κλάδο του αυτοκινήτου τα έτη 2010 και 2011, αναλύοντας ξεχωριστά τις πωλήσεις του κάθε τύπου καυσίμου, υποθέτοντας ότι τα δεδομένα είναι απλής διαστρωμάτωσης.

- **Μελέτη της πορείας των πωλήσεων των βενζινοκίνητων οχημάτων**

Θεωρώντας ότι τα δεδομένα έχουν απλώς διαστρωματική μορφή και αγνοώντας την ύπαρξη δεδομένων που αφορούν τις πωλήσεις των πετρελαιοκίνητων οχημάτων, η εκτίμηση των επιπτώσεων της οικονομικής κρίσης στην πορεία των πωλήσεων των βενζινοκίνητων οχημάτων θα πραγματοποιηθεί μέσω του παρακάτω υποδείγματος:

$$gasoline_sales_i = \beta_0 + \beta_1 year11 + u \quad (4.1)$$

όπου η εξαρτημένη μεταβλητή $gasoline_sales_i$ περιγράφει το πλήθος των βενζινοκίνητων οχημάτων που πωλήθηκαν σε κάθε χώρα, ενώ η ανεξάρτητη μεταβλητή $year11$, είναι μία ψευδομεταβλητή, η οποία λαμβάνει την τιμή 0, αν οι πωλήσεις αφορούν την περίοδο του 2010 και την τιμή 1, αν οι πωλήσεις αφορούν την περίοδο του 2011. Η μορφή του συνόλου των δεδομένων, καθώς και οι εκτιμήσεις των συντελεστών του υποδείγματος της παλινδρόμησης παρουσιάζονται στον Πίνακα 4.5 και στον Πίνακα 4.6 αντίστοιχα.

Πίνακας 4.5

Σύνολο δεδομένων πωλήσεων βενζινοκίνητων οχημάτων

I		Unleades_Sales	year11	i	Unleades_Sales	year11
1	Αυστρία	161.982	0	19	161.690	1
2	Βέλγιο	131.909	0	20	141.336	1
3	Δανία	83.398	0	21	90.474	1
4	Φινλανδία	62.797	0	22	70.279	1
5	Γαλλία	657.487	0	23	608.367	1
6	Γερμανία	1.694.346	0	24	1.678.852	1
7	Ελλάδα	135.841	0	25	87.912	1
8	Ιρλανδία	2.432	0	26	2.907	1
9	Ιταλία	33.433	0	27	26.963	1
10	Λουξεμβούργο	1.061.214	0	28	783.883	1
11	Ολλανδία	12.332	0	29	11.622	1
12	Πορτογαλία	386.054	0	30	398.540	1
13	Ισπανία	32.066	0	31	33.618	1
14	Σουηδία	73.520	0	32	46.635	1
15	Ην. Βασίλειο	288.712	0	33	239.991	1
16	Ισλανδία	142.235	0	34	117.724	1
17	Νορβηγία	204.132	0	35	213.237	1
18	Ελβετία	1.094.626	0	36	958.979	1

Πίνακας 4.6

Εκτίμηση συντελεστών του υποδείγματος 4.1

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	347695,333	107771,107		3,226	,003
year11	-32528,167	152411,362	-,037	-,213	,832

Όπως γίνεται εμφανές από τον Πίνακα 4.6, το 2010 πωλήθηκαν κατά μέσο όρο 347.695 βενζινοκίνητα οχήματα, ενώ το 2011 οι μέσες πωλήσεις αυτών μειώθηκαν κατά 32.528. Παρόλα αυτά η μείωση αυτή δεν φέρεται να είναι στατιστικά σημαντική και ως εκ τούτου το υπόδειγμα 4.1 δεν περιγράφει με ικανοποιητικό τρόπο τη μεταβολή αυτή.

- **Μελέτη της πορείας των πωλήσεων των πετρελαιοκίνητων οχημάτων**

Αντίστοιχα για την κατηγορία των πετρελαιοκίνητων οχημάτων, η εκτίμηση των επιπτώσεων της οικονομικής κρίσης στην πορεία των πωλήσεων τους, θα πραγματοποιηθεί μέσω του υποδείγματος:

$$diesel_sales_i = \beta_0 + \beta_1 year11 + u \quad (4.2)$$

όπου η εξαρτημένη μεταβλητή $diesel_sales_i$ περιγράφει το πλήθος των πετρελαιοκίνητων οχημάτων που πωλήθηκαν σε κάθε χώρα. Η μορφή του συνόλου των δεδομένων, καθώς και οι εκτιμήσεις των συντελεστών του υποδείγματος της παλινδρόμησης παρουσιάζονται στον Πίνακα 4.7 και στον Πίνακα 4.8 αντίστοιχα.

Συνοψίζοντας τα αποτελέσματα του υποδείγματος 4.2 προκύπτει ότι το 2010 πωλήθηκαν κατά μέσο όρο 373.299 πετρελαιοκίνητα οχήματα ενώ το 2011 οι πωλήσεις αυτών αυξήθηκαν κατά 23.093. Παρόλα αυτά η αύξηση αυτή δεν φέρεται να είναι στατιστικά σημαντική και ως εκ τούτου το υπόδειγμα 4.2 δεν περιγράφει με ικανοποιητικό τρόπο τη μεταβολή αυτή.

Ο λόγος που τα υποδείγματα 4.1 και 4.2 δεν περιγράφουν με ικανοποιητικό τρόπο την όποια μεταβολή των πωλήσεων στα βενζινοκίνητα ή στα πετρελαιοκίνητα οχήματα, οφείλεται κυρίως στο γεγονός ότι παρόλο που τα δεδομένα είναι μορφής πάνελ, δεν έχουν συμπεριληφθεί τυχόν απαραίτητες επιδράσεις. Το πρόβλημα αυτό αντιμετωπίζεται μέσω του υποδείγματος σταθερών ή απαραίτητων επιδράσεων.

Πίνακας 4.7

Σύνολο δεδομένων πωλήσεων πετρελαιοκίνητων οχημάτων

I		Diesel_Sales	year11	i	Diesel_Sales	year11
1	Αυστρία	166.581	0	19	194.455	1
2	Βέλγιο	415.431	0	20	430.875	1
3	Δανία	70.189	0	21	79.270	1
4	Φινλανδία	44.549	0	22	50.892	1
5	Γαλλία	1.594.182	0	23	1.595.862	1
6	Γερμανία	1.221.913	0	24	1.494.782	1
7	Ελλάδα	5.660	0	25	9.768	1
8	Ιρλανδία	674	0	26	2.131	1
9	Ιταλία	55.013	0	27	62.915	1
10	Λουξεμβούργο	900.365	0	28	965.856	1
11	Ολλανδία	37.394	0	29	38.259	1
12	Πορτογαλία	96.513	0	30	157.304	1
13	Ισπανία	95.688	0	31	104.727	1
14	Σουηδία	149.944	0	32	106.769	1
15	Ην. Βασίλειο	693.303	0	33	568.060	1
16	Ισλανδία	147.449	0	34	187.260	1
17	Νορβηγία	88.321	0	35	103.609	1
18	Ελβετία	936.220	0	36	982.274	1

Πίνακας 4.8

Εκτίμηση συντελεστών του υποδείγματος 4.2

Model	Unstandardized Coefficients		Standardized Coefficients	T	Sig.
	B	Std. Error	Beta		
(Constant)	373299,389	117714,076		3,171	,003
year11	23093,278	166472,843	,024	,139	,890

- **Μέθοδος των Διαφορών σε Διαφορές**

Θεωρώντας και σε αυτό το σημείο ότι τα δεδομένα δεν είναι μορφής πάνελ, αλλά ομαδοποιημένης διαστρωμάτωσης, θα εξεταστεί η μεταβολή των πωλήσεων των πετρελαιοκίνητων οχημάτων με τη μέθοδο των Διαφορών σε Διαφορές που αναπτύχθηκε εκτενώς στις παραγράφους του 2^{ου} Κεφαλαίου. Πιο συγκεκριμένα, ως ομάδα ελέγχου ορίζονται οι πωλήσεις των βενζινοκίνητων οχημάτων, ως ομάδα αντιμετώπισης οι πωλήσεις των πετρελαιοκίνητων οχημάτων, ως χρονική περίοδος προ της χρηματοπιστωτικής κρίσης στην Ευρωπαϊκή Ένωση το έτος 2010 και ως χρονική

περίοδος μετά την χρηματοπιστωτική κρίση το 2011. Το υπόδειγμα που χρησιμοποιείται στην παρούσα μέθοδο είναι της μορφής:

$$sales_i = \beta_0 + \beta_1 year11 + \beta_2 diesel + \beta_3 year11 \cdot diesel + u \quad (4.3)$$

όπου η εξαρτημένη μεταβλητή $sales_i$ περιγράφει τις πωλήσεις της κάθε διαστρωμάτωσης, η ψευδομεταβλητή $year11$ λαμβάνει την τιμή 0, αν η παρατήρηση προέρχεται από το έτος 2010 και την τιμή 1, αν η παρατήρηση προέρχεται από το έτος 2011, η ψευδομεταβλητή $diesel$ λαμβάνει την τιμή 0, αν η παρατήρηση προέρχεται από το σύνολο των βενζινοκίνητων οχημάτων και την τιμή 1, αν προέρχεται από το σύνολο των πετρελαιοκίνητων οχημάτων, ενώ τέλος, η μεταβλητή $year11 \cdot diesel$ περιγράφει την αλληλεπίδραση των δύο ψευδομεταβλητών. Ο συντελεστής β_3 του υποδείματος 4.3, είναι ο συντελεστής των διαφορών σε διαφορές και περιγράφει τον τρόπο που μεταβλήθηκαν οι πωλήσεις των πετρελαιοκίνητων οχημάτων εξαιτίας της χρηματοπιστωτικής κρίσης. Από την ανάλυση των δεδομένων προκύπτουν οι παρακάτω συντελεστές:

Πίνακας 4.9

Εκτίμηση συντελεστών του υποδείματος 4.3

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
(Constant)	347.695,333	112852,149		3,081	,003
year11	-32.528,167	159597,040	-,035	-,204	,839
diesel	25.604,056	159597,040	,027	,160	,873
year11·diesel	55.621,444	225704,299	,052	,246	,806

Συνοψίζοντας τα αποτελέσματα του πίνακα 4.9 προκύπτουν τα παρακάτω συμπεράσματα:

- Το έτος 2010 οι μέσες πωλήσεις των βενζινοκίνητων οχημάτων είναι ίσες με 347.695 οχήματα.
- Το έτος 2011 οι μέσες πωλήσεις των βενζινοκίνητων οχημάτων είναι ίσες με $347.695 - 32.528 = 315.167$ οχήματα.
- Το έτος 2010 οι μέσες πωλήσεις των πετρελαιοκίνητων οχημάτων είναι ίσες με $347.695 + 25.604 = 373.299$ οχήματα.

- Το έτος 2011 οι μέσες πωλήσεις των πετρελαιοκίνητων οχημάτων που διαμορφώθηκαν λόγω της χρηματοπιστωτικής κρίσης, είναι ίσες με $347.695 - 32.528 + 25.604 + 55621 = 396.392$ οχήματα.

Παρά την εκτίμηση των συντελεστών μέσω της μεθόδου των Διαφορών σε Διαφορές, είναι εμφανές από τον Πίνακα 4.9, ότι οι συντελεστές αυτοί δεν είναι στατιστικά σημαντικοί. Αυτό οφείλεται στο γεγονός ότι, ενώ τα δεδομένα είναι μορφής πάνελ, έχουν αναλυθεί σαν ομαδοποιημένα διαστρωματικά. Με αυτό τον τρόπο δε συμπεριλαμβάνονται στο υπόδειγμα 4.3 οι απαραίτητες επιδράσεις της κάθε χώρας, με αποτέλεσμα οι εκτιμήσεις να είναι μεροληπτικές. Η καταλληλότερη μέθοδος για την εκτίμηση των επιπτώσεων της χρηματοπιστωτικής κρίσης στις πωλήσεις των πετρελαιοκίνητων οχημάτων, είναι αυτή της μεθόδου των διαφορών ή των σταθερών επιδράσεων για δεδομένα πάνελ, έτσι ώστε να απαληφθούν οι όποιες απαραίτητες επιδράσεις. Στην παρούσα μελέτη λαμβάνει χώρα η μέθοδος των διαφορών, χωρίς αυτό να σημαίνει πως δε μπορεί να πραγματοποιηθεί και η μέθοδος των σταθερών επιδράσεων.

Έλεγχος Παράλληλης Τάσης

Όπως αναφέρθηκε στην Παράγραφο 2.6, η παράλληλη τάση αποτελεί την πιο σημαντική προϋπόθεση για την εκτίμηση ενός υποδείγματος μέσω της μεθόδου των διαφορών σε διαφορές. Στην παρούσα μελέτη, ο έλεγχος της παράλληλης τάσης, πραγματοποιείται μέσα από στατιστικούς ελέγχους, για το αν τα δείγματα κατανέμονται με όμοιο τρόπο. Πιο συγκεκριμένα, ελέγχεται, για τα έτη 2003 έως και 2010, αν οι πωλήσεις των βενζινοκίνητων και των πετρελαιοκίνητων οχημάτων για κάθε χρονική περίοδο χωριστά, κατανέμονται με όμοιο τρόπο, με χρήση του ελέγχου t-test και του μη παραμετρικού ελέγχου των Mann-Whitney για δύο ανεξάρτητα δείγματα, αναλόγως της κατανομής των δειγμάτων. Παράλληλα, στην περίπτωση που τα δεδομένα είναι μορφής πάνελ και όχι απλώς ομαδοποιημένα διαστρωματικά, θα πρέπει να δειχθεί, ότι στην ομάδα ελέγχου δεν υπήρξαν σημαντικές μεταβολές στις δύο χρονικές περιόδους. Το παραπάνω ελέγχεται με τον έλεγχο Signed Rank του Wilcoxon για δύο εξαρτημένα δείγματα.

- **Έλεγχος Κανονικότητας των Δειγμάτων**

Πριν από κάθε έλεγχο όμοιας κατανομής ανεξάρτητων ή εξαρτημένων δειγμάτων, πρέπει να ελεγχθεί αν το κάθε δείγμα από την κάθε χρονική περίοδο, κατανέμεται κανονικά έτσι ώστε να πραγματοποιηθεί στη συνέχεια ο κατάλληλος παραμετρικός ή μη παραμετρικός έλεγχος. Στον επόμενο συγκεντρωτικό πίνακα, παρουσιάζονται τα αποτελέσματα των ελέγχων περί ύπαρξης κανονικότητας των Kolmogorov-Smirnov, με αρχική υπόθεση την κανονικότητα του κάθε δείγματος. Ο έλεγχος πραγματοποιείται σε επίπεδο σημαντικότητας 5%.

Αποτελέσματα κανονικότητας κατά Kolmogorov-Smirnov

	Τύπος Καυσίμου	Significant Value	Απόφαση
2003	Gasoline	0,095	Μη απόρριψη
	Diesel	0,081	Μη απόρριψη
2004	Gasoline	0,162	Μη απόρριψη
	Diesel	0,072	Μη απόρριψη
2005	Gasoline	1,139	Μη απόρριψη
	Diesel	0,063	Μη απόρριψη
2006	Gasoline	0,144	Μη απόρριψη
	Diesel	0,047	Απόρριψη σε ε.σ 5%
2007	Gasoline	0,189	Μη απόρριψη
	Diesel	0,029	Απόρριψη σε ε.σ 5%
2008	Gasoline	0,165	Μη απόρριψη
	Diesel	0,029	Απόρριψη σε ε.σ 5%
2009	Gasoline	0,033	Απόρριψη σε ε.σ 5%
	Diesel	0,040	Απόρριψη σε ε.σ 5%
2010	Gasoline	0,106	Μη απόρριψη
	Diesel	0,038	Απόρριψη σε ε.σ 5%

Όπως προκύπτει από τα αποτελέσματα των ελέγχων, οι πωλήσεις βενζινοκίνητων και πετρελαιοκίνητων οχημάτων για τα έτη 2003, 2004 και 2005 κατανέμονται κανονικά και ως εκ τούτου, για αυτές τις ομάδες θα πραγματοποιηθεί έλεγχος μέσω t test. Για τα υπόλοιπα έτη, απορρίπτεται η υπόθεση κανονικότητας ενός εκ των δύο ή και των

δύο δειγμάτων και επομένως, θα πραγματοποιηθεί ο έλεγχος U των Mann-Whitney, για κάθε ζεύγος δειγμάτων, με αρχική υπόθεση την όμοια κατανομή αυτών.

- **Έλεγχος Όμοιας Κατανομής των Δειγμάτων**

Στο συγκεντρωτικό πίνακα που ακολουθεί, παρουσιάζονται τα αποτελέσματα των ελέγχων t και U για το αν τα δείγματα κατανέμονται με όμοιο τρόπο. Όπως προκύπτει από τα αποτελέσματα των ελέγχων, είναι εμφανής η επικράτηση της αρχικής υπόθεσης και ως εκ τούτου οι πωλήσεις των βενζινοκίνητων και των πετρελαιοκίνητων οχημάτων κατανέμονται με όμοιο τρόπο σε όλες τις χρονικές περιόδους. Συμπερασματικά, η υπόθεση της παράλληλης τάσης ισχύει και ως εκ τούτου, αν δεν υπήρχε απότομη μεταβολή στο οικονομικό περιβάλλον, οι πωλήσεις των οχημάτων με πετρελαιοκινητήρα, θα ακολουθούσαν την ίδια πορεία με αυτή των βενζινοκινητήρων.

Αποτελέσματα όμοιας κατανομής δειγμάτων

	Τύπος Καυσίμου	Τύπος Ελέγχου	Sig. Value	Απόφαση
2003	Gasoline VS Diesel	t test	0,588	Μη απόρριψη
2004	Gasoline VS Diesel	t test	0,878	Μη απόρριψη
2005	Gasoline VS Diesel	t test	0,950	Μη απόρριψη
2006	Gasoline VS Diesel	U test	0,372	Μη απόρριψη
2007	Gasoline VS Diesel	U test	0,696	Μη απόρριψη
2008	Gasoline VS Diesel	U test	0,767	Μη απόρριψη
2009	Gasoline VS Diesel	U test	0,791	Μη απόρριψη
2010	Gasoline VS Diesel	U test	1,000	Μη απόρριψη

- **Έλεγχος Μεταβολής Ομάδας Ελέγχου**

Στον επόμενο συγκεντρωτικό πίνακα, ακολουθούν τα αποτελέσματα του ελέγχου των Kolmogorov-Smirnov για το αν τα δείγματα των πωλήσεων των βενζινοκίνητων οχημάτων κατανέμονται κανονικά για τα έτη 2010 και 2011.

Αποτελέσματα κανονικότητας κατά Kolmogorov-Smirnov

	Τύπος Καυσίμου	Significant Value	Απόφαση
2010	Gasoline	0,106	Μη απόρριψη
2011	Gasoline	0,097	Μη απόρριψη

Από αυτόν προκύπτει ότι οι πωλήσεις των βενζινοκίνητων οχημάτων, κατανέμονται κανονικά. Επομένως, ο έλεγχος της ύπαρξης σημαντικής μεταβολής στα δύο δείγματα, θα πραγματοποιηθεί μέσω του ελέγχου t για δύο εξαρτημένα δείγματα, με αρχική υπόθεση τη μη μεταβολή των δειγμάτων. Τα αποτελέσματα του ελέγχου εμφανίζονται στον επόμενο πίνακα και γίνεται εμφανές το γεγονός, ότι οι πωλήσεις των βενζινοκίνητων οχημάτων δε μεταβλήθηκαν σημαντικά στα έτη 2010 και 2011.

Έλεγχος μεταβολής της ομάδας ελέγχου

	Τύπος Καυσίμου	Sig. Value	Απόφαση
2010 VS 2011	Gasoline	0,068	Μη απόρριψη

Μέθοδος Διαφορών με Χρήση Δεδομένων Πάνελ

Όπως αναφέρθηκε στο Κεφάλαιο 2, οι επιπτώσεις ενός συμβάντος δύνανται να εκτιμηθούν είτε μέσω του υποδείγματος (2.8), που αφορά ομαδοποιημένα διαστρωματικά δεδομένα, είτε μέσω του υποδείγματος (2.10), στην περίπτωση που τα δεδομένα είναι μορφής πάνελ. Στην παρούσα μελέτη, λόγω της φύσης των δεδομένων, χρησιμοποιείται η μεθοδολογία του υποδείγματος (2.10), με ομάδα αντιμετώπισης τα πετρελαιοκίνητα οχήματα και ομάδα ελέγχου τα βενζινοκίνητα, υποθέτοντας ότι η οικονομική κρίση έχει κάνει την εμφάνισή της στα νοικοκυριά και στη συνείδηση του κόσμου, στο τέλος του 2010.

Η εκτίμηση των επιπτώσεων της οικονομικής κρίσης στον κλάδο του αυτοκινήτου, λαμβάνει χώρα μέσω του υποδείγματος παρατηρήτων επιδράσεων:

$$sales_{it} = \beta_0 + \beta_1 year11_t + \beta_2 diesel_{it} + \alpha_i + u_{it}, \quad (4.4)$$

όπου η εξαρτημένη μεταβλητή $sales_{it}$, περιγράφει το πλήθος των οχημάτων που πωλήθηκαν σε κάθε χώρα, τα έτη 2010 και 2011, η μεταβλητή $year11_t$ είναι η χρονική ψευδομεταβλητή που λαμβάνει την τιμή 0, αν η παρατήρηση προέρχεται από το έτος 2010 και την τιμή 1, αν προέρχεται από το έτος 2011, η μεταβλητή $diesel_{it}$ είναι η

ψευδομεταβλητή που δηλώνει σε ποια ομάδα ανήκει η παρατήρηση και λαμβάνει την τιμή 0, αν η παρατήρηση προέρχεται από την ομάδα ελέγχου των βενζινοκίνητων οχημάτων και την τιμή 1, αν προέρχεται από την ομάδα αντιμετώπισης των πετρελαιοκίνητων οχημάτων και τέλος, η μεταβλητή α_i περιγράφει τις παρατηρήσιμες επιδράσεις του υποδείγματος, οι οποίες μπορεί να περιλαμβάνουν την κουλτούρα και τη φιλοσοφία του κάθε κράτους ως προς την αγορά πετρελαιοκίνητου οχήματος, τη νομοθεσία, το μέσο ετήσιο εισόδημα και ούτω καθεξής.

Η απαλοιφή των παρατηρήσιμων επιδράσεων, πραγματοποιείται μέσω της λήψης των διαφορών και ως εκ τούτου προκύπτει το υπόδειγμα:

$$\Delta sales_i = \beta_1 + \beta_2 \Delta diesel_i + \Delta u_{it}, \quad (4.5)$$

στο οποίο μελετάται η μεταβολή των πωλήσεων στα βενζινοκίνητα και πετρελαιοκίνητα οχήματα, στα έτη 2010 και 2011. Η μέση συνολική μεταβολή των πωλήσεων των πετρελαιοκίνητων είναι ίση με:

$$E(\Delta sales | \Delta diesel_i = 1) = \beta_1 + \beta_2 \quad (4.6)$$

ενώ των βενζινοκίνητων οχημάτων είναι ίση με:

$$E(\Delta sales | \Delta diesel_i = 0) = \beta_1 \quad (4.7)$$

Αφαιρώντας τις σχέσεις (4.6) και (4.7) προκύπτει η εκτίμηση των επιπτώσεων της οικονομικής κρίσης στις πωλήσεις των πετρελαιοκίνητων οχημάτων, η οποία είναι ίση με το συντελεστή β_2 . Ως εκ τούτου:

$$E(\Delta y | \Delta d_i = 1) - E(\Delta y | \Delta d_i = 0) = \beta_2$$

Στον Πίνακα 4.10 παρουσιάζεται το σύνολο των δεδομένων μετά τη λήψη των διαφορών, ενώ στον Πίνακα 4.11, παρουσιάζονται τα αποτελέσματα του υποδείγματος 4.5.

Πίνακας 4.10

Σύνολο δεδομένων μεταβολής πωλήσεων οχημάτων

I		Δsales	Δyear11	Δdiesel	i	Δsales	Δyear11	Δdiesel
1	Αυστρία	-292	1	0	19	27.874	1	1
2	Βέλγιο	9.427	1	0	20	15.444	1	1
3	Δανία	7.076	1	0	21	9.081	1	1
4	Φινλανδία	7.482	1	0	22	6.343	1	1
5	Γαλλία	-49.120	1	0	23	1.680	1	1
6	Γερμανία	-15.494	1	0	24	272.869	1	1
7	Ελλάδα	-47.929	1	0	25	4.108	1	1
8	Ιρλανδία	-6.469	1	0	26	7.901	1	1
9	Ιταλία	-277.331	1	0	27	65.491	1	1
10	Λουξεμβούργο	-710	1	0	28	865	1	1
11	Ολλανδία	12.487	1	0	29	60.790	1	1
12	Πορτογαλία	-26.885	1	0	30	-43.175	1	1
13	Ισπανία	-48.721	1	0	31	-125.243	1	1
14	Σουηδία	-24.511	1	0	32	39.811	1	1
15	Ην. Βασίλειο	-135.647	1	0	33	46.054	1	1
16	Ισλανδία	475	1	0	34	1.457	1	1
17	Νορβηγία	1.552	1	0	35	9.039	1	1
18	Ελβετία	9.105	1	0	36	15.288	1	1

Πίνακας 4.11

Εκτίμηση συντελεστών του υποδείγματος 4.4

Model		Unstandardized Coefficients		Standardized	T	Sig.
		B	Std. Error	Beta		
1	Beta ₁	-32528,056	17198,879	-,426	-1,891	,067
	Beta ₂	55621,222	24322,888	,515	2,287	,029

Από τον Πίνακα 4.11 είναι εμφανές, ότι η μέση μεταβολή των πωλήσεων των πετρελαιοκίνητων οχημάτων για τα έτη 2010 και 2011, είναι ίση με: $E(\Delta sales | \Delta diesel_i = 1) = \beta_1 + \beta_2 = -32.528 + 55.621 = 23.093$ μονάδες προϊόντος, ενώ η μέση μεταβολή των βενζινοκίνητων οχημάτων για τα έτη 2010 και 2011 είναι ίση με: $E(\Delta sales | \Delta diesel_i = 0) = \beta_1 = -32.528$ μονάδες προϊόντος. Η μεταβολή στις πωλήσεις των πετρελαιοκίνητων οχημάτων που οφείλεται στην οικονομική κρίση, είναι ίση με το συντελεστή β_2 του υποδείγματος, ο οποίος είναι ο συντελεστής των διαφορών σε διαφορές. Πιο συγκεκριμένα, η μεταβολή αυτή είναι ίση

με 55.621 οχήματα και είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 5% διότι $sig_{beta_2} = 2,9\% < 5\%$, όπως στατιστικά σημαντική εμφανίζεται και η μεταβολή των βενζινοκίνητων οχημάτων σε επίπεδο σημαντικότητας 10% λόγω του ότι $sig_{beta_1} = 6,7\% < 10\%$.

4.5 Ανακεφαλαίωση

Στο παρόν κεφάλαιο, παρουσιάστηκαν τα εμπειρικά αποτελέσματα της επίδρασης της οικονομικής κρίσης, στις καταναλωτικές προτιμήσεις στον κλάδο του αυτοκινήτου, μέσω της μεθόδου των διαφορών σε διαφορές με χρήση δεδομένων πάνελ, για τις χώρες Αυστρία, Βέλγιο, Δανία, Φινλανδία, Γαλλία, Γερμανία, Ελλάδα, Ιρλανδία, Ιταλία, Λουξεμβούργο, Ολλανδία, Πορτογαλία, Ισπανία, Σουηδία, Ηνωμένο Βασίλειο που ανήκουν στην Ευρωπαϊκή Ένωση των 15 χωρών (*EU15*) και τις Ισλανδία, Νορβηγία και Ελβετία που ανήκουν στην Ευρωπαϊκή Ζώνη Ελεύθερων Συναλλαγών (*EFTA3*). Αρχικά, παρουσιάστηκε η ιστορική αναδρομή της Ευρωπαϊκής οικονομικής κρίσης και εν συνεχεία, έγινε η παρουσίαση των δεδομένων για το σύνολο των χωρών που ανήκουν στην *EU15* και στην *EFTA3*. Ύστερα, πραγματοποιήθηκαν απλές εμπειρικές μελέτες, χωρίς να συμπεριληφθούν οι απαραίτητες επιδράσεις των δεδομένων πάνελ, με στόχο να αναδειχθεί η αδυναμία τους, καθώς και στατιστικοί έλεγχοι για την απόδειξη και διασφάλιση της παράλληλης τάσης των δειγμάτων. Τέλος, έλαβε χώρα η μέθοδος των διαφορών σε διαφορές, όπου προέκυψε στατικά σημαντική θετική μεταβολή στην αγορά πετρελαιοκίνητων οχημάτων εξαιτίας της οικονομικής κρίσης και στατιστικά σημαντική αρνητική μεταβολή στην αγορά βενζινοκίνητων οχημάτων.

Πανεπιστήμιο Πειραιώς

ΠΑΡΑΡΤΗΜΑ

Αποτελέσματα κανονικότητας των δειγμάτων κατά Kolmogorov-Smirnov

➤ 2003 Gasoline

One-Sample Kolmogorov-Smirnov Test		Gasoline2003
N		18
Normal Parameters ^{a,b}	Mean	444488,78
	Std. Deviation	603135,474
	Absolute	,291
Most Extreme Differences	Positive	,291
	Negative	-,235
Kolmogorov-Smirnov Z		1,234
Asymp. Sig. (2-tailed)		,095

a. Test distribution is Normal.

b. Calculated from data.

➤ 2003 Diesel

One-Sample Kolmogorov-Smirnov Test		Diesel2003
N		18
Normal Parameters ^{a,b}	Mean	345103,94
	Std. Deviation	480814,600
	Absolute	,298
Most Extreme Differences	Positive	,298
	Negative	-,237
Kolmogorov-Smirnov Z		1,266
Asymp. Sig. (2-tailed)		,081

a. Test distribution is Normal.

b. Calculated from data.

➤ **2004 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2004
N		18
Normal Parameters ^{a,b}	Mean	417503,61
	Std. Deviation	551531,829
	Absolute	,264
Most Extreme Differences	Positive	,264
	Negative	-,230
Kolmogorov-Smirnov Z		1,121
Asymp. Sig. (2-tailed)		,162

a. Test distribution is Normal.

b. Calculated from data.

➤ **2004 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2004
N		18
Normal Parameters ^{a,b}	Mean	389410,28
	Std. Deviation	536032,588
	Absolute	,304
Most Extreme Differences	Positive	,304
	Negative	-,235
Kolmogorov-Smirnov Z		1,290
Asymp. Sig. (2-tailed)		,072

a. Test distribution is Normal.

b. Calculated from data.

➤ **2005 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2005
N		18
Normal Parameters ^{a,b}	Mean	408540,44
	Std. Deviation	540688,368
	Absolute	,272
Most Extreme Differences	Positive	,272
	Negative	-,232
Kolmogorov-Smirnov Z		1,155
Asymp. Sig. (2-tailed)		,139

a. Test distribution is Normal.

b. Calculated from data.

➤ **2005 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2005
N		18
Normal Parameters ^{a,b}	Mean	397279,50
	Std. Deviation	538668,993
	Absolute	,310
Most Extreme Differences	Positive	,310
	Negative	-,232
Kolmogorov-Smirnov Z		1,316
Asymp. Sig. (2-tailed)		,063

a. Test distribution is Normal.

b. Calculated from data.

➤ **2006 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2006
N		18
Normal Parameters ^{a,b}	Mean	402352,00
	Std. Deviation	533147,839
	Absolute	,270
Most Extreme Differences	Positive	,270
	Negative	-,232
Kolmogorov-Smirnov Z		1,147
Asymp. Sig. (2-tailed)		,144

a. Test distribution is Normal.

b. Calculated from data.

➤ **2006 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2006
N		18
Normal Parameters ^{a,b}	Mean	417815,39
	Std. Deviation	562117,335
	Absolute	,323
Most Extreme Differences	Positive	,323
	Negative	-,231
Kolmogorov-Smirnov Z		1,370
Asymp. Sig. (2-tailed)		,047

a. Test distribution is Normal.

b. Calculated from data.

➤ **2007 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2007
N		18
Normal Parameters ^{a,b}	Mean	383686,06
	Std. Deviation	497244,770
	Absolute	,256
Most Extreme Differences	Positive	,256
	Negative	-,227
Kolmogorov-Smirnov Z		1,085
Asymp. Sig. (2-tailed)		,189

a. Test distribution is Normal.

b. Calculated from data.

➤ **2007 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2007
N		18
Normal Parameters ^{a,b}	Mean	438183,00
	Std. Deviation	573124,959
	Absolute	,343
Most Extreme Differences	Positive	,343
	Negative	-,225
Kolmogorov-Smirnov Z		1,455
Asymp. Sig. (2-tailed)		,029

a. Test distribution is Normal.

b. Calculated from data.

➤ **2008 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2008
N		18
Normal Parameters ^{a,b}	Mean	356728,28
	Std. Deviation	481843,965
	Absolute	,263
Most Extreme Differences	Positive	,263
	Negative	-,233
Kolmogorov-Smirnov Z		1,116
Asymp. Sig. (2-tailed)		,165

a. Test distribution is Normal.

b. Calculated from data.

➤ **2008 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2008
N		18
Normal Parameters ^{a,b}	Mean	396667,94
	Std. Deviation	516127,464
	Absolute	,343
Most Extreme Differences	Positive	,343
	Negative	-,223
Kolmogorov-Smirnov Z		1,456
Asymp. Sig. (2-tailed)		,029

a. Test distribution is Normal.

b. Calculated from data.

➤ **2009 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2009
N		18
Normal Parameters ^{a,b}	Mean	410112,00
	Std. Deviation	670942,469
	Absolute	,337
Most Extreme Differences	Positive	,337
	Negative	-,271
Kolmogorov-Smirnov Z		1,432
Asymp. Sig. (2-tailed)		,033

a. Test distribution is Normal.

b. Calculated from data.

➤ **2009 Diesel**

One-Sample Kolmogorov-Smirnov Test

		Diesel2009
N		18
Normal Parameters ^{a,b}	Mean	349099,50
	Std. Deviation	481977,305
	Absolute	,330
Most Extreme Differences	Positive	,330
	Negative	-,235
Kolmogorov-Smirnov Z		1,400
Asymp. Sig. (2-tailed)		,040

a. Test distribution is Normal.

b. Calculated from data.

➤ **2010 Gasoline**

One-Sample Kolmogorov-Smirnov Test

		Gasoline2010
N		18
Normal Parameters ^{a,b}	Mean	347695,33
	Std. Deviation	474429,300
	Absolute	,286
Most Extreme Differences	Positive	,286
	Negative	-,233
Kolmogorov-Smirnov Z		1,212
Asymp. Sig. (2-tailed)		,106

a. Test distribution is Normal.

b. Calculated from data.

➤ **2010 Diesel**

One-Sample Kolmogorov-Smirnov Test

		DieselSales201
		0
N		18
Normal Parameters ^{a,b}	Mean	373299,39
	Std. Deviation	483624,450
	Absolute	,332
Most Extreme Differences	Positive	,332
	Negative	-,221
Kolmogorov-Smirnov Z		1,409
Asymp. Sig. (2-tailed)		,038

a. Test distribution is Normal.

b. Calculated from data.

➤ **Gasoline 2011**

One-Sample Kolmogorov-Smirnov Test

		GasolineSales2 011
N		18
Normal Parameters ^{a,b}	Mean	315167,17
	Std. Deviation	439366,426
	Absolute	,290
Most Extreme Differences	Positive	,290
	Negative	-,239
Kolmogorov-Smirnov Z		1,231
Asymp. Sig. (2-tailed)		,097

a. Test distribution is Normal.

b. Calculated from data.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ελληνική

- Jeffrey M. Wooldridge (2013). *Εισαγωγή στην οικονομετρία*, Εκδόσεις Παπαζήση
- Γιάννης Βαρουφάκης, Τάσος Πατώκος, Λευτέρης Τσερκέζης, Χρήστος Κουτσοπέτρος (2011). *Η οικονομική κρίση στην Ελλάδα και την Ευρώπη το 2011*, Παρατηρητήριο Οικονομικών και Κοινωνικών Εξελίξεων

Ξένα

- Jeffrey M. Wooldridge (2012). *Introductory Econometrics: A modern approach*
- Joshua D. Angrist, Jörn-Steffen Pischke (2008). *Mostly Harmless Econometrics: An Empiricist's Companion*
- Michael R. Roberts, Toni M. Whited (2011). *Endogeneity in Empirical Corporate Finance*
- Chris Brooks (2008). *Introductory Econometrics for Finance*
- Jeffrey M. Wooldridge (2001). *Econometric Analysis of Cross Section and Panel Data*
- Badi H. Baltagi (2005). *Econometric Analysis of Panel Data*
- Bruce D. Meyer (1995). *Natural and Quasi-Experiments in Econometrics*, Technical Working Paper Series
- David Card, Alan B. Krueger (1994). *Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania*, The American Economy Review
- Bruce D. Meyer, W. Kip Viscusi, David L. Durbin (1995). *Workers' Compensation and Injury Duration: Evidence from a Natural Experiment*
- John Kennan (1995). *The Elusive Effects of Minimum Wages*
- Damodar N. Gujarati (2004). *Basic Econometrics*

Διαδίκτυο

<http://epp.eurostat.ec.europa.eu/>

<http://www.acea.be/>

www.ineobservatory.gr

Πανεπιστήμιο Πειραιώς

Πανεπιστήμιο Πειραιώς