

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



**ΠΜΣ «Πληροφοριακά Συστήματα & Υπηρεσίες»
Ειδίκευση: «Προηγμένα Πληροφοριακά Συστήματα»**

**Σύγκριση συστάδων Ευρωπαϊκών χωρών μέσω
ανάλυσης χρονοσειρών και παλινδρομήσεων για τις
επιδράσεις της διαφθοράς στις Άμεσες Ξένες
Επενδύσεις**

Τσιροζίδης Δημήτριος (ΜΕ2153)

Επιβλέπων Καθηγητής : Φιλιππάκης Μιχαήλ

Πειραιάς, Φεβρουάριος 2023

Περίληψη

Η παρούσα διπλωματική εργασία μελετά τις επιδράσεις της διαφθοράς στις άμεσες ξένες επενδύσεις χρησιμοποιώντας τόσο συμβατικούς δείκτες όσο και εναλλακτικές μετρήσεις για αυτήν. Η εργασία ξεκινά με ανάλυση χρονοσειρών και στην συνέχεια επικεντρώνεται σε ένα πλαίσιο παλινδρομήσεως μέσω της γλώσσας προγραμματισμού R. Η ανάλυση βασίζεται σε δύο συστάδες χωρών, αυτή της ανατολικής Ευρώπης και αυτή του υπολοίπου Ευρώπης. Κατά τη διάρκεια της ανάλυσης μελετώνται επίσης άλλοι παράγοντες που μπορεί να επηρεάζουν τους ξένους επενδυτές στην λήψη αποφάσεων. Τα αποτελέσματα δείχνουν πως οι μέχρι τώρα συμβατικές αναλύσεις είναι πολύ δύσκολο να δώσουν απάντηση στο εάν και πως η διαφθορά επηρεάζει τις εירוές άμεσων επενδύσεων. Αντίθετα η εργασία έδειξε πως εναλλακτικά proxies ίσως είναι σε θέση να βοηθήσουν, ειδικά όταν λαμβάνονται υπόψιν παράγοντες όπως οι ιδιαιτερότητες κάθε χώρας ή ακόμα και η χώρα προέλευσης του επενδυτή

Λέξεις κλειδιά: Χρονοσειρές, ARIMA, Ανάλυση Παλινδρόμησης, Άμεσες Ξένες Επενδύσεις, Διαφθορά, Βιομηχανία

Ευχαριστίες - Acknowledgements

I would like to thank Professor Michael Filippakis for his valuable guidance and for the trust he showed me during the preparation of my dissertation. I would also like to thank Dr. Poulou Marilena for the assistance in supervising the analysis of the data, her contribution to the experimental part of the dissertation and her helpful comments in the research analysis.

Πίνακας Περιεχομένων

Κεφάλαιο 1.Εισαγωγή.....	1
Κεφάλαιο 2. Ανασκόπηση Βιβλιογραφίας.....	5
2.1 Θεωρία.....	5
2.2 Επιδράσεις της διαφθοράς στις εισροές άμεσων ξένων επενδύσεων.....	7
2.2.1 Ευρήματα αρνητικής σχέσης.....	8
2.2.2 Ευρήματα θετικής σχέσης.....	9
2.2.3 Ευρήματα για μη ύπαρξη σχέσης.....	10
2.2.4 Μέθοδοι – Υποθέσεις.....	10
Κεφάλαιο 3.Αναδρομή στη Θεωρία.....	12
3.1 Ανάλυση Δεδομένων.....	12
3.2 Ανάλυση Παλινδρόμησης.....	17
3.2.1 Παλινδρόμηση και Συσχέτιση.....	17
3.2.2 Απλή Γραμμική Παλινδρόμηση.....	18
3.2.3 Υποθέσεις απλού γραμμικού μοντέλου.....	21
3.2.4 Έλεγχος υποθέσεων απλού γραμμικού μοντέλου.....	21
3.3 Τύποι Δεδομένων.....	25
3.3.1 Time-Series Data.....	26
3.3.2 Cross-Sectional Data.....	27
3.3.3 Panel Data.....	28
3.3.4 Continuous και Discreet data - Συνεχή και Διακριτά.....	29
3.3.5 Cardinal, Ordinal και Nominal αριθμοί.....	30

3.4 Θεωρία χρονοσειράς.....	30
3.4.1 Ορισμός χρονοσειράς.....	30
3.4.2 Βασικά Χαρακτηριστικά Χρονοσειράς.....	33
3.4.3 Συσχέτιση σε χρονοσειρά.....	35
3.4.3.1 Συσχέτιση σε χρονοσειρά-Λευκός θόρυβος –White noise.....	35
3.4.3.2 Συσχέτιση σε χρονοσειρά-Τυχαίος περίπατος–Random walk.....	36
3.4.4 Γραμμική συσχέτιση.....	37
3.4.5 Θεωρία Μοντέλων ARIMA.....	38
3.4.6 Μοντέλα Αυτοπαλινδρόμησης - Autoregressive models- AR(p).....	42
3.4.7 Μοντέλα Κινητού Μέσου Όρου - Moving Average MA(q).....	43
3.4.8 Μοντέλα ARIMA (p,d,q).....	43
3.4.9 Πρόβλεψη με μοντέλα ARIMA.....	45
Κεφάλαιο 4.Ανάλυση Μεταβλητών και Μεθοδολογιών.....	46
4.1 Εξαρτημένη Μεταβλητή.....	46
4.2 Κύρια Ανεξάρτητη Μεταβλητή.....	46
4.3 Εξαρτημένες Μεταβλητές Ελέγχου.....	47
4.4 Κύριο Συμβατικό Μοντέλο.....	50
Κεφάλαιο 5.Εναλλακτική Προσέγγιση: Διεφθαρμένοι Κλάδοι και Επενδύσεις.....	51
5.1 Περιγραφή.....	51
5.2 Νέες μεταβλητές.....	53
5.3 Μεθοδολογία.....	54
Κεφάλαιο 6. Αποτελέσματα Ανάλυσης Χρονοσειρών.....	55
Κεφάλαιο 7. Αποτελέσματα Παλινδρομήσεων.....	68

7.1 Συμβατικό Μοντέλο.....	68
7.2 Προσέγγιση home-host country.....	86
7.3 Εναλλακτική προσέγγιση.....	93
Κεφάλαιο 8.Συμπεράσματα.....	103
Κεφάλαιο 9.Βιβλιογραφία.....	107
Κεφάλαιο 10.Arpendix.....	110

Πίνακες

Πίνακας 1: Αρχικό Μοντέλο.....	69
Πίνακας 2: Αρχικό Μοντέλο χωρίς ετεροσκεδαστικότητα.....	71
Πίνακας 3α: Εισαγωγή μεταβλητών ελέγχου σχετικών με την πολιτική σταθερότητα και τις κοινωνικές ελευθερίες για το υπόλοιπο της Ευρώπης.....	73
Πίνακας 3β: Εισαγωγή μεταβλητών ελέγχου σχετικών με την πολιτική σταθερότητα και τις κοινωνικές ελευθερίες για την ανατολική Ευρώπη.....	76
Πίνακας 4α: Εισαγωγή μεταβλητών ελέγχου σχετικών με δείκτες χρημ. επιδόσεων και το ανθρώπινο κεφάλαιο για το υπόλοιπο της Ευρώπης.....	79
Πίνακας 4β: Εισαγωγή μεταβλητών ελέγχου σχετικών με δείκτες χρημ.επιδόσεων και το ανθρώπινο κεφάλαιο για την ανατολική Ευρώπη.....	82
Πίνακας 5α: Πλήρες μοντέλο για το υπόλοιπο της Ευρώπης.....	84
Πίνακας 5β: Πλήρες μοντέλο για την ανατολική Ευρώπη.....	85
Πίνακας 6: Πλήρη μοντέλα για τις δύο ομάδες χωρών.....	91
Πίνακας 7α: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς - εξαρτημένη μεταβλητή δείκτης Cpi για το υπόλοιπο της Ευρώπης.....	95
Πίνακας 7β: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς - εξαρτημένη μεταβλητή δείκτης Cpi για την ανατολική Ευρώπη.....	96
Πίνακας 8α: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή την FDI για το υπόλοιπο της Ευρώπης.....	100
Πίνακας 8β: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή την FDI για την ανατολική Ευρώπη.....	101

Γραφήματα

Γράφημα 1: Γραφική απεικόνιση των δύο συστάδων.....	5
Γράφημα 2: Τα 5 βήματα ανάλυσης δεδομένων.....	13
Γράφημα 3: Ετήσιοι μέσοι όροι συστάδων δείκτη CPI για έτη 2003-2020.....	55
Γράφημα 4: Χρονοσειρά υπολοίπου Ευρώπης.....	59
Γράφημα 5: Αυτοσυσχετίσεις υπολοίπου Ευρώπης.....	61
Γράφημα 6: ARIMA 3 ετών για υπόλοιπο Ευρώπης.....	63
Γράφημα 7: Χρονοσειρά ανατολικής Ευρώπης.....	64
Γράφημα 8: Αυτοσυσχετίσεις ανατολικής Ευρώπης.....	66
Γράφημα 9: ARIMA 3 ετών ανατολικής Ευρώπης.....	67
Γράφημα 10: Χώρες με υψηλότερα δημόσια χρέη στην Ευρώπη (%ΑΕΠ).....	80
Γράφημα 11: Υποσυστάδα Ρωσίας.....	88
Γράφημα 12: Υποσυστάδα Γερμανίας.....	89
Γράφημα 13: Μέσος όρος μεταβλητής ROL για κάθε συστάδα.....	99

Εικόνες

Εικόνα 1: Οι 4 πιο διεφθαρμένοι τομείς σύμφωνα με την έρευνα.....	52
Εικόνα 2: Δείγμα δεδομένων στην R.....	58
Εικόνα 3: DF test για υπόλοιπο Ευρώπης.....	60
Εικόνα 4: KPSS test για υπόλοιπο Ευρώπης.....	60
Εικόνα 5: AIC υπολοίπου Ευρώπης.....	62
Εικόνα 6: DF test για ανατολική Ευρώπη.....	65
Εικόνα 7: KPSS test για ανατολική Ευρώπη.....	65
Εικόνα 8: AIC ανατολικής Ευρώπης.....	67

1.Εισαγωγή

Ένας από τους τομείς, ο οποίος απασχολεί έντονα τους επιστήμονες και γενικότερα τους ερευνητές εδώ και δεκαετίες είναι η ανάλυση της λήψης αποφάσεων των επενδυτών. Πιο συγκεκριμένα, οι παράγοντες οι οποίοι οδηγούν τους επενδυτές στο να κατανέμουν τα χρήματα και τους πόρους τους σε μία εταιρεία ή στη περίπτωση της παρούσας εργασίας, σε μία χώρα, έχουν αποτελέσει αντικείμενο εκτεταμένης μελέτης. Η ερώτηση όμως αυτή δεν φαίνεται να έχει μία σαφή απάντηση στη βιβλιογραφία, καθώς η ερμηνεία της επενδυτικής συμπεριφοράς είναι ιδιαίτερα περίπλοκη και μπορεί να διαφέρει κατά περίπτωση.

Σύμφωνα με την Ευρωπαϊκή Τράπεζα Επενδύσεων (EIB) (FDI and productivity convergence in central and eastern Europe, 2009), οι άμεσες ξένες επενδύσεις – FDI (Foreign Direct Investment) είναι σε θέση να αυξήσουν τους ρυθμούς παραγωγικότητας σε πληθώρα τομέων στην οικονομία μίας χώρας. Αυτή η στόχευση στην παραγωγικότητα των άμεσων ξένων επενδύσεων, είναι κάτι που πρέπει να αναφερθεί, καθώς άλλοι τύποι χρηματικών ροών όπως σε εμβάσματα ή σε μορφή ξένης βοήθειας (foreign aid) στοχεύουν περισσότερο τομείς που σχετίζονται με την κατανάλωση και όχι τόσο σε τομείς που μπορούν να ενισχύσουν τους ρυθμούς οικονομικής ανάπτυξης μίας χώρας (Angelon, 2018). Σύμφωνα με τα Ηνωμένα Έθνη και συγκεκριμένα τον φορέα Nation’s Millennium Development Goals (MDGs), οι κυβερνήσεις των αναπτυσσόμενων χωρών πρέπει να συνεχίσουν να ακολουθούν στρατηγικές και να διαμορφώνουν πολιτικές οι οποίες έλκουν ξένες άμεσες επενδύσεις.

Πάρα την πολυπλοκότητα της ερμηνείας των επιλογών των επενδυτών, έχουν προκύψει κάποια ερευνητικά συμπεράσματα όσον αφορά τις επιλογές των ξένων επενδυτών τα οποία είναι αρκετά ισχυρά και επαναλαμβανόμενα. Αυτό που είναι σαφές στη βιβλιογραφία είναι ότι υπάρχει μία πληθώρα παραγόντων που μπορεί να επηρεάσουν την απόφαση για κατανομή των πόρων τους. Πιο συγκεκριμένα, οι άμεσοι επενδυτές κατά κύριο λόγο προτιμούν σταθερούς και δημοκρατικούς θεσμούς (Ahlquist, 2006). Μεταξύ άλλων σημαντικοί παράγοντες για την αύξηση των

FDI εισροών είναι το πόσο ανοιχτή είναι μία χώρα στις επενδύσεις, το νομικό της σύστημα αλλά και ο αρχικός βαθμός ανάπτυξης της.

Όπως γίνεται αντιληπτό, ορισμένοι παράγοντες έχουν αναλυθεί και αναγνωρισθεί στην χρηματοοικονομική βιβλιογραφία, όμως ο παράγοντας που έχει απασχολήσει περισσότερο τη βιβλιογραφία, χωρίς όμως να παραχθούν σαφή αποτελέσματα, είναι η διαφθορά και το κατά πόσο αυτή επηρεάζει ή όχι, θετικά ή αρνητικά, την συμπεριφορά και τις επιλογές των επενδυτών. Οι επιδράσεις λοιπόν της διαφθοράς παραμένουν σύμφωνα με τη βιβλιογραφία ένα από τα εμπόδια που σχετίζονται με την οικονομική ανάπτυξη φτωχότερων χωρών, αφού έχει ήδη αποδειχθεί ότι οι άμεσες ξένες επενδύσεις μπορούν να λειτουργήσουν θετικά όσον αφορά τους ρυθμούς ανάπτυξης, αν και κάποια αποτελέσματα διαφέρουν σε ορισμένες αναπτυσσόμενες χώρες (Lensink και Morrissey, 2006). Τα τελευταία χρόνια, εκτεταμένη έρευνα έχει πραγματοποιηθεί για την κατανόηση εκείνων των παραγόντων που φέρνουν και ενισχύουν την ανάπτυξη ως ένα εργαλείο μείωσης της φτώχειας και για αυτό το λόγο οι επιδράσεις των άμεσων ξένων επενδύσεων βρέθηκαν στο προσκήνιο. Όπως ανακοινώθηκε από την εργασία της Ευρωπαϊκής Τράπεζας Επενδύσεων (EIB) (Impact of FDI on economic growth, 2020), οι εταιρείες σε χώρες μεσαίου εισοδήματος είναι πιο πιθανό να απορροφήσουν τα πλεονεκτήματα που συνοδεύουν τις ξένες επενδύσεις και άρα πιο πιθανό να αυξήσουν την παραγωγικότητα τους. Όμως, σύμφωνα και πάλι με την ΕΤΕ, διαφθορά και χαμηλή ποιότητα θεσμών σε αυτές τις χώρες μπορεί να επιφέρει αντιπαραγωγικά αποτελέσματα. Επειδή αυτό το θέμα συνδέεται στενά με τα οικονομικά ανάπτυξης τα περισσότερα άρθρα-εργασίες στοχεύουν τις χώρες της υποσαχάριας Αφρικής και της νοτιοανατολικής Ασίας, δύο περιοχές γνωστές για τα μεγάλα ποσοστά του πληθυσμού που ζουν κάτω από τα όρια της φτώχειας.

Η παρούσα διπλωματική εργασία, λαμβάνοντας ως αφορμή τα υψηλά ποσοστά διαφθοράς στην ανατολική Ευρώπη και την προσπάθεια των χωρών της να αναπτυχθούν ανάλογα με τις χώρες της δυτικής Ευρώπης, θα βασιστεί στη δημιουργία 2 συστάδων-ομάδων χωρών. Πιο συγκεκριμένα θα δημιουργηθεί μία συστάδα 19 χωρών ανατολικής Ευρώπης και μία συστάδα 19 χωρών από την δυτική Ευρώπη, έχοντας ως στόχο την εξαγωγή συμπερασμάτων μέσω της σύγκρισης των

αποτελεσμάτων. Αυτό που προκαλεί εντύπωση είναι πως παρά την υψηλή διαφθορά της ανατολικής Ευρώπης και τον μεγάλο αριθμό χωρών και ανθρώπων σε αυτή φαίνεται σαν περιοχή να έχει σχεδόν “αγνοηθεί” στην βιβλιογραφία. Ίσως οι ερευνητές να αποθαρρύνονται από τις μεγάλες πολιτισμικές διαφορές που διέπουν την ανατολικές χώρες αλλά και τα διαφορετικά επίπεδα οικονομικού φιλελευθερισμού (Costica, 2012). Αυτές οι διαφορές όμως εάν αντιμετωπιστούν σωστά μπορούν αντιθέτως να λειτουργήσουν ευεργετικά όσον αφορά τα συμπεράσματα για τις επιδράσεις της διαφθοράς. Πέρα όμως από τα προβλήματα που δημιουργούνται από τοπικές διαφορές και ιδιαιτερότητες, πολλές έρευνες αντιμετωπίζουν οικονομετρικά προβλήματα όπως η ενδογένεια, ετεροσκεδαστικότητα, ελλειπούσες μεταβλητές κ.α. Σύμφωνα με την Ευρωπαϊκή Τράπεζα Επενδύσεων οι χώρες διαφέρουν αρκετά μεταξύ τους ακόμα και στις επενδυτικές τους ανάγκες. Για παράδειγμα αν πάρουμε τη συστάδα της ανατολικής Ευρώπης, χώρες όπως η Τσεχία και η Ουγγαρία έχουν μεγαλύτερη ανάγκη για επενδύσεις στους τομείς των υποδομών και της εκπαίδευσης. Αντίστοιχα, χώρες όπως της Βαλτικής χρειάζονται ξένα κεφάλαια κυρίως στον τομέα της ενέργειας. Η τράπεζα κατέληξε επίσης ότι η ανατολική Ευρώπη χρειάζεται άμεσα ξένες επενδύσεις στον ιδιωτικό τομέα, στις μεταφορές, στον τουρισμό αλλά και σε πολλούς άλλους τομείς. Αυτός είναι και ένας από τους λόγους που η παρούσα διπλωματική εργασία θα επικεντρωθεί στις ροές άμεσων ξένων επενδύσεων και όχι σε άλλα είδη χρηματικών ροών.

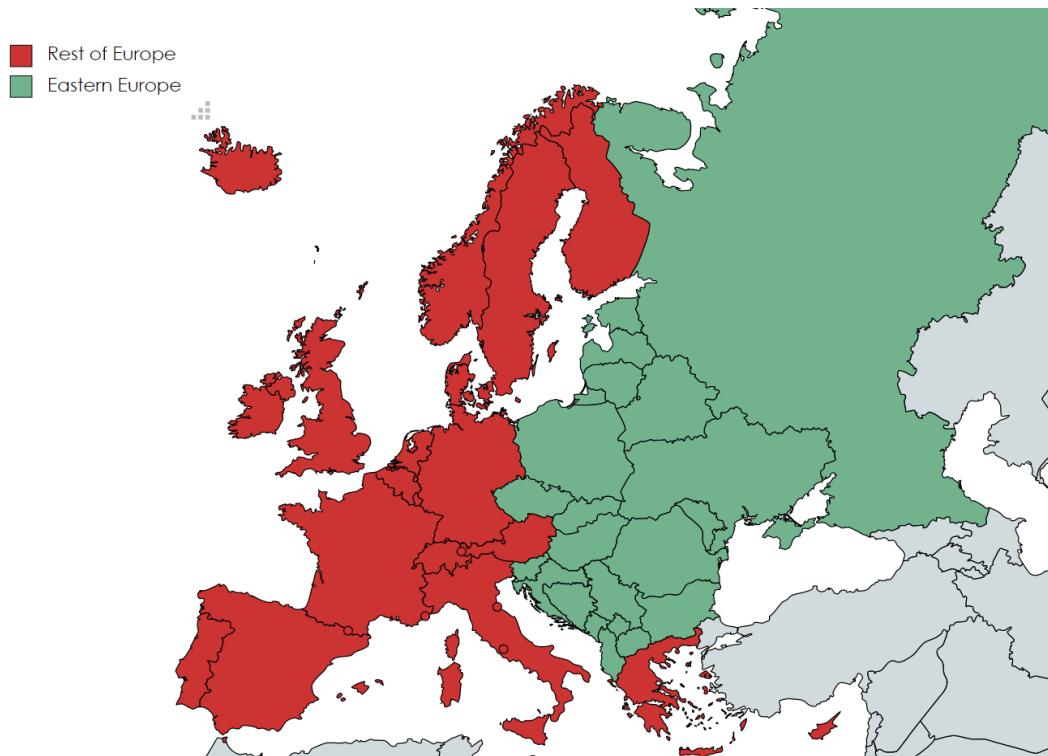
Οι ξένοι επενδυτές με τα κεφάλαια τους μπορούν να οδηγήσουν σε αυξημένες παραγωγικότητες και βελτιωμένες τεχνολογίες πράγματα απαραίτητα για τις εγχώριες επιχειρήσεις και άρα για τις χώρες στην προσπάθεια τους να φτάσουν τις οικονομικές επιδόσεις των χωρών της δεύτερης συστάδας, δηλαδή εκείνες της δυτικής Ευρώπης. Παρότι η μετάβαση από τα παλαιότερα συγκεντρωτικά καθεστώτα της ανατολικής Ευρώπης σε πιο φιλελεύθερα συστήματα έχει ξεκινήσει εδώ και δεκαετίες και κάποιες από αυτές τις χώρες της ανατολικής συστάδας θεωρούνται πλέον μεσαίου εισοδήματος, υπάρχουν κάποιες που συμπεριλαμβάνονται ακόμα στις αναπτυσσόμενες χώρες και οικονομίες, σύμφωνα με το Διεθνές Νομισματικό Ταμείο. Οι χώρες αυτές λοιπόν έχουν κάποιες ομοιότητες μεταξύ τους, σε επίπεδο

δομών, μεγέθους, ανθρώπινου κεφαλαίου, αλλά έχουν και μεγάλες διαφορές όπως την ποιότητα των θεσμών αλλά και το πόσο ανοιχτές είναι σε ξένες επενδύσεις. Οι δύο τελευταίοι παράγοντες είναι αρκετά σημαντικοί σύμφωνα με τη βιβλιογραφία για τους επενδυτές και την κατανομή των πόρων τους. Η βασικότερη διαφορά ανάμεσα στις δύο συστάδες είναι τα επίπεδα διαφθοράς των χωρών που τις απαρτίζουν. Σύμφωνα με τον οργανισμό Transparency International οι χώρες της ανατολικής Ευρώπης έχουν 15 μονάδες χαμηλότερη βαθμολογία από τις χώρες της υπόλοιπης Ευρώπης κατά μέσο όρο. Ενδεικτικό της διαφοράς είναι πως το 2018 η μέση βαθμολογία των χωρών της ανατολικής Ευρώπης ήταν 55, ενώ της Ολλανδίας μόλις 18. Πέρα όμως από τις άμεσες ξένες επενδύσεις, υπάρχουν στοιχεία που δείχνουν πως η διαφθορά επηρεάζει της εγχώριες επιχειρήσεις σε ότι έχει να κάνει με την λήψη αποφάσεων τους (Blagojević και Damijan, 2013).

Καταλήγοντας, η εργασία θα βασιστεί πάνω σε δύο συστάδες (οι χώρες των συστάδων αναφέρονται στο Appendix). Η πρώτη αφορά την ανατολική Ευρώπη, μία περιοχή με μεγάλη ανάγκη για επενδύσεις και τα πλεονεκτήματα που αυτές φέρνουν στο παρόν αλλά και στο μέλλον. Η δεύτερη θα αφορά χώρες της λεγόμενης δυτικής Ευρώπης, όχι τόσο με τη γεωγραφική έννοια αλλά περισσότερο με πολιτισμική, πολιτική και ιστορική προσέγγιση. Συγκεκριμένα θα ερευνηθεί μέσα σε ένα πλαίσιο μοντέλων παλινδρόμησης το εάν οι άμεσες ξένες επενδύσεις ανά συστάδα επηρεάζονται από τα επίπεδα διαφθοράς και δεύτερον ποιοι άλλοι παράγοντες μπορεί να επηρεάζουν τις επενδύσεις με βάση τα δεδομένα τα οποία θα συλλεγούν.

Η δεύτερη ενότητα πραγματοποιεί μία αναδρομή στη βιβλιογραφία που είναι σχετική με τις άμεσες ξένες επενδύσεις και τη διαφθορά αλλά και τις υποθέσεις που θα εξεταστούν στην παρούσα εργασία. Στην ενότητα 3 αναλύονται θεωρητικές έννοιες σχετικές με όσα χρησιμοποιήθηκαν στο πρακτικό κομμάτι της εργασίας. Στις ενότητες 4 και 5 περιγράφονται οι μεταβλητές και οι μεθοδολογίες, ενώ στις ενότητες 6 και 7 παρουσιάζονται τα αποτελέσματα της ανάλυσης χρονοσειρών και παλινδρομήσεων αντίστοιχα. Τέλος, στην ενότητα 8 αναφέρονται τα συμπεράσματα.

Γράφημα 1: Γραφική απεικόνιση των δύο συστάδων



2.Ανασκόπηση Βιβλιογραφίας

Σε αυτό το κεφάλαιο διερευνάται η θεωρία και τα αποτελέσματα της σχέσης ανάμεσα στην διαφθορά και στις εισροές άμεσων ξένων επενδύσεων. Στην υποενότητα 2.1 παρουσιάζονται διάφοροι ορισμοί της διαφθοράς καθώς και διαφορετικές και αντικρουόμενες απόψεις αναφορικά με την αναμενόμενη επίδραση της διαφθοράς στις εισροές άμεσων ξένων επενδύσεων. Στην υποενότητα 2.2 παρουσιάζονται ευρήματα από τη βιβλιογραφία που αναφέρονται στη σχέση των δύο κύριων αντικειμένων της εργασίας, της διαφθοράς και των εισροών άμεσων ξένων επενδύσεων.

2.1 Θεωρία

Ο αριθμός των παραγόντων που μπορούν να θεωρηθούν σαν εν δυνάμει κίνητρα για πιθανές εισροές άμεσων ξένων επενδύσεων είναι μεγάλος, και οι σχετικές απόψεις στη διεθνή βιβλιογραφία είναι αντικρουόμενες, για διάφορους λόγους.

Για να μελετηθεί η σχέση διαφθοράς και ξένων επενδύσεων είναι αρχικά απαραίτητο να συζητηθεί ο ορισμός της πρώτης. Στην παρούσα εργασία η διαφθορά ορίζεται σαν η κατάχρηση της δημόσιας εξουσίας για ίδιο όφελος. Ο ορισμός αυτός δεν απέχει πολύ από τον τρόπο που ο Macrae (1982) περιέγραψε τη διαφθορά, δηλαδή σαν μια διευθέτηση ή συναλλαγή ανάμεσα σε δύο πλευρές που επηρεάζει την κατανομή πόρων, άμεσα ή μελλοντικά και καταχράται την δημόσια ή συλλογική ευθύνη με σκοπό το κέρδος.

Η πιο κοινή άποψη αναφορικά με τη διαφθορά είναι ότι πρόκειται για ανέντιμη συμπεριφορά, κυρίως από άτομα που βρίσκονται σε θέση ισχύος όπως κυβερνητικούς αξιωματούχους ή μεγαλοστελεχη του επιχειρηματικού τομέα. Μπορεί να περιλαμβάνει συναλλαγές «κάτω από το τραπέζι», ξέπλυμα χρήματος, δωροδοκίες ή ακόμα και χειραγώγηση εκλογών. Αυτός είναι και ο ορισμός σύμφωνα με τον οργανισμό Transparency International. Το κίνητρο για την συμπεριφορά αυτή είναι το κέρδος, ενώ αγνοούνται οι αρνητικές επιδράσεις και η επιβάρυνση που προκαλείται στην κοινωνία. Στους υποστηρικτές της προσέγγισης αυτής περιλαμβάνονται οι Globerman και Shapiro (2002), Wei (2000) και οι Simon και Eitzen (1990). Η άποψη αυτή βλέπει την διαφθορά σαν αιτία δημιουργίας αρνητικών συνεπειών κυρίως στην οικονομία μιας χώρας και κατά συνέπεια στην καθημερινότητα των πολιτών της. Οι υποστηρικτές της προσέγγισης αυτής πιστεύουν ότι η διαφθορά επιβαρύνει και τους επενδυτές και ότι φαινόμενα όπως νεποτισμός, δωροδοκίες κλπ. μπορούν να μεγαλώσουν το πρόβλημα αβεβαιότητας εκείνων που δεσμεύουν τους πόρους τους περιμένοντας συγκεκριμένα μελλοντικά αποτελέσματα. Συνεπώς η αβεβαιότητα αυτή οδηγεί σε μείωση των ξένων επενδύσεων σε διεφθαρμένες χώρες, αφού οι επενδυτές αξιολογούν αρνητικά αυτό το επιχειρηματικό περιβάλλον και βλέπουν τη διαφθορά μόνο σαν πρόβλημα και εμπόδιο. Αν συνεπώς η άποψη αυτή είναι η κοντινότερη στην πραγματικότητα θα πρέπει να αναμένουμε αρνητική συσχέτιση ανάμεσα στα επίπεδα διαφθοράς και στις εισροές άμεσων ξένων επενδύσεων.

Υπάρχει όμως και η αντίθετη άποψη, ότι δηλαδή η διαφθορά μπορεί να επιδρά θετικά στις εισροές άμεσων ξένων επενδύσεων. Κάποιοι υποψήφιοι επενδυτές μπορεί να σκεφθούν πως η αυξημένη διαφθορά μιας χώρας τους προσφέρει τη δυνατότητα να παρακάμψουν τυχόν νόμους ή κανονισμούς που θα τους δημιουργούσαν προβλήματα και τελικά να πετύχουν τους στόχους τους ταχύτερα και ευκολότερα (Lui, 1985). Αυτό έχει παρατηρηθεί όχι μόνο σε φτωχές αλλά και στις μεσαίου εισοδήματος χώρες, με αιτία κυρίως την ανεπάρκεια και βραδύτητα των θεσμών και της Δικαιοσύνης. Ενας σημαντικός αριθμός μεσαίου εισοδήματος χωρών συναντώνται στην Ανατολική Ευρώπη και θα είχε ενδιαφέρον να εξεταστούν. Οι υποστηρικτές της προσέγγισης αυτής βλέπουν τη διαφθορά σαν κομμάτι των επιχειρηματικών δαπανών και ισχυρίζονται ότι η ταχύτητα με την οποία προχωρούν τα επιχειρηματικά σχέδια, μαζί με άλλα πλεονεκτήματα που παρέχει στους επενδυτές η διαφθορά υπερκαλύπτουν κατά πολύ το αυξημένου κόστους που αυτή τους δημιουργεί (Rose-Ackerman, 1978). Κάποιες μελέτες, όπως αυτές των Leff (1964), Leyes (1965) και Bayley (1966), παρουσίασαν κάποιους θετικούς εξωτερικούς παράγοντες που συσχετίζονται με την αυξημένη διαφθορά. Για παράδειγμα, χώρες με μεγάλο πρόβλημα γραφειοκρατίας, που συνήθως απωθεί τους επενδυτές, έχουν την ευκαιρία λόγω της διαφθοράς να κερδίσουν επενδύσεις από επενδυτές που είναι διατεθειμένοι να κινηθούν σε ένα τέτοιο περιβάλλον. Ο Haggard, (1990) επισήμανε πως ένα διεφθαρμένο περιβάλλον μπορεί να είναι θελκτικό για επένδυτες που βλέπουν δυνατότητα να φοροδιαφύγουν, καθώς και να ξεπεράσουν τυχόν δυσκολίες που πηγάζουν από την εργασιακή νομοθεσία. Οι μελέτες αυτές επικεντρώνονται στα οφέλη των εισροών άμεσων ξένων επενδύσεων στις χώρες του τρίτου κόσμου. Αν και στην ανατολική Ευρώπη η ποιότητα ζωής είναι πολύ καλύτερη, σε κάποιες χώρες τα επίπεδα φιλελευθεροποίησης και συμμόρφωσης με τα ευρωπαϊκά πρότυπα είναι χαμηλά και τα γραφειοκρατικά προβλήματα εξακολουθούν να υπάρχουν.

2.2 Επιδράσεις της διαφθοράς στις εισροές άμεσων ξένων επενδύσεων.

Η σημασία της διαδικασίας λήψης αποφάσεων εκ μέρους των επενδυτών οδήγησε πολλούς ερευνητές να εξετάσουν τους παράγοντες εκείνους που θα αντιμετώπιζονταν από τους επενδυτές σαν θετικά κίνητρα για να διαθέσουν πόρους

σε ένα συγκεκριμένο πρόγραμμα. Οι καθοριστικοί παράγοντες για τις επιλογές σχετικά με ξένες επενδύσεις είναι ανάμεσα στα ζητήματα που πάντα συγκέντρωναν το ενδιαφέρον των οικονομολόγων. Αναφορικά με τις άμεσες ξένες επενδύσεις η απελευθέρωση του εμπορίου αναφέρεται συχνά σαν κρίσιμος παράγοντας για μεγαλύτερες ροές επενδύσεων (Sebastian-Andrei Labes, 2015). Οι Hermes και Lensink (2003) υποστηρίζουν επίσης ότι το επίπεδο ανάπτυξης του χρηματοπιστωτικού συστήματος στην υπό εξέταση χώρα είναι σημαντικό στοιχείο για τις άμεσες ξένες επενδύσεις. Η ποιότητα των θεσμών και η πολιτική σταθερότητα φαίνεται επίσης να εκτιμούνται από τους ξένους επενδυτές (Ahlquist, 2006), μαζί με τη γεωγραφική απόσταση και την ανάπτυξη του ανθρωπίνου κεφαλαίου που επίσης θεωρούνται σημαντικοί παράγοντες για υψηλές εισροές άμεσων ξένων επενδύσεων (Hattari και Rajan, 2011). Σε ένα άρθρο που προκάλεσε αίσθηση οι Borensztein et al. (1998) ισχυρίζονται πως οι άμεσες ξένες επενδύσεις και η αποτελεσματικότητά τους συσχετίζονται έντονα με το απόθεμα ανθρωπίνου κεφαλαίου της αντίστοιχης χώρας. Οι ερευνητές φαίνονται να συγκλίνουν στην άποψη πως η σχέση ανάμεσα στις εισροές άμεσων ξένων επενδύσεων και στις μεταβλητές που αναφέρθηκαν παραπάνω είναι θετική. Από την άλλη πλευρά, τα αποτελέσματα της έρευνας αναφορικά με την διαφθορά στην σχετική βιβλιογραφία είναι ανάμεικτα. Ειδικότερα, κάποιες μελέτες έχουν επισημάνει αρνητική συσχέτιση αναμεσα στην εισροή άμεσων ξένων επενδύσεων και στην εκτεταμένη διαφθορά στην υπό εξέταση χώρα, κάποιες άλλες θετική, ενώ άλλες δεν βρήκαν σημαντικά στοιχεία που να ενισχύουν την άποψη ότι υπάρχει συσχέτιση.

2.2.1 Ευρήματα αρνητικής σχέσης

Ακολουθώντας την άποψη ότι η διαφθορά επηρεάζει αρνητικά την εισροή άμεσων ξένων επενδύσεων, οι Freckleton et al (2012) στην ερευνά τους σχετικά με τις εισροές άμεσων επενδύσεων, την ανάπτυξη και το ρόλο της διαφθοράς επισήμαναν πως ένα χαμηλότερο επίπεδο διαφθοράς μπορεί να βελτιώσει το αποτέλεσμα των άμεσων ξένων επενδύσεων στην ανάπτυξη των αναπτυσσόμενων χωρών. Συνεπώς η έρευνα τους περιγράφει αρνητική σχέση ανάμεσα στη διαφθορά και στην εισροή άμεσων

ξένων επενδύσεων. Αντίστοιχα, ο Wei (2000), εξετάζοντας τη σχέση ανάμεσα στη διαφθορά και στην διμερή (bilateral) εισροή άμεσων ξένων επενδύσεων σε δείγμα 57 χωρών με χρήση ποικιλίας δεικτών διαφθοράς, συμπεραίνει πως η διαφθορά έχει αρνητική επίδραση στις εισροές άμεσων ξένων επενδύσεων. Επιπρόσθετα ο συγγραφέας επισημαίνει ότι η επίδραση της διαφθοράς στην εισροή άμεσων ξένων επενδύσεων είναι πολύ μεγαλύτερη όταν συγκρίνεται με χρηματικές ροές άλλης μορφής. Ο Wei (1997) ακόμα συμπεραίνει πως η αβεβαιότητα που δημιουργείται από την διαφθορά έχει επίσης αρνητική επίδραση στις εισροές άμεσων ξένων επενδύσεων. Ο Bailey (2018), χρησιμοποιώντας μία μετα-αναλυτική Hedges και Olkin διαδικασία βασιζόμενη σε 97 προηγούμενες μελέτες που εξέταζαν θεσμικούς παράγοντες και ελκυστικότητα στην εισροή άμεσων ξένων επενδύσεων βρίσκει επίσης αρνητική σχέση μεταξύ των τελευταίων και της διαφθοράς. Ο Mauro (1995) καταλήγει στο ότι ένα υψηλό επίπεδο διαφθοράς επιδρά αρνητικά στον όγκο των επενδύσεων και έχει συνεπώς αρνητικό αποτέλεσμα στην οικονομική ανάπτυξη μίας χώρας. Τέλος οι Habib & Zurawicki (2002) στην 89-σέλιδη ανάλυσή τους συμπεραίνουν πως η επίδραση της διαφθοράς στην εισροή άμεσων ξένων επενδύσεων είναι σημαντικά αρνητική και ότι οι ξένοι επενδυτές προτιμούν να μην διακινδυνεύουν τους πόρους τους σε ένα περιβάλλον διαφθοράς.

2.2.2 Ευρήματα θετικής σχέσης

Από την άλλη πλευρά υπάρχει ένας σημαντικός αριθμός μελετών που βρίσκει θετική σχέση ανάμεσα στην διαφθορά και στις ξένες επενδύσεις. Ο Aidt (2003) ισχυρίζεται ότι η διαφθορά, με τη μορφή δωροδοκιών, μπορεί να αυξήσει την αποτελεσματικότητα, ιδίως σε χώρες όπου βασιλεύει η γραφειοκρατία. Αντίστοιχα ο Bardhan (1997) δηλώνει ότι η διαφθορά μπορεί να προσφέρει τη δυνατότητα παράκαμψης τυχόν εχθρικών κυβερνητικών αποφάσεων και της γραφειοκρατίας. Οι Bellos και Subasat (2012) εξετάζοντας έναν αριθμό από αναπτυσσόμενες οικονομίες παρουσιάζουν στατιστικά στοιχεία που δείχνουν θετική επίδραση της διαφθοράς στην εισροή άμεσων ξένων επενδύσεων. Ο Leff (1964) ακολουθεί μια κάπως διαφορετική προσέγγιση. Στην μελέτη του, η διαφθορά αποτελεί μέθοδο για τη μείωση της αβεβαιότητας. Μέσα από διεφθαρμένες πρακτικές όπως δωροδοκίες, οι

επενδυτές μπορούν να εξασφαλίσουν πληροφορίες και να ανοίξουν πόρτες που αλλιώς θα παρέμεναν κλειστές. Έτσι οι επενδύσεις τους θα είναι πιο ασφαλείς. Κατά συνέπεια, η διαφθορά είναι δυνατό να ενθαρρύνει αυτό το είδος επενδυτών να κατευθύνουν τους πόρους τους σε χώρες με πρόβλημα διαφθοράς, αυξάνοντας τις ροές των άμεσων ξένων επενδύσεων.

2.2.3 Ευρήματα για μη ύπαρξη σχέσης

Υπάρχει ένας σημαντικός αριθμός πρόσφατων μελετών που δεν κατάφεραν να βρουν σημαντική επίδραση της διαφθοράς στις εισροές άμεσων ξένων επενδύσεων. Ειδικότερα, οι Abed and Davoodi (2000), με χρήση μιας ανάλυση που χωρίστηκε σε ένα cross-sectional και μετά σε ένα panel κομμάτι κατέληξαν αρχικά ότι χώρες με χαμηλά επίπεδα διαφθοράς προτιμώνται από τους ξένους επενδυτές, αφού οι χώρες αυτές έλκουν μεγαλύτερες εισροές άμεσων ξένων επενδύσεων κατά κεφαλή. Σε επόμενα κεφάλαια της ίδιας μελέτης όμως το εύρημα αυτό αμφισβητείται, αφού μετά από την χρήση μεταβλητών ελέγχου σχετικές με την ποιότητα των δομών κάθε χώρας τα προαναφερθέντα αποτελέσματα εμφανίζονται στατιστικά ασήμαντα. Οι Stein και Daude (2001), δεν κατάφεραν να επισημάνουν καμιά σημαντική σχέση ανάμεσα στη διαφθορά και στις εισροές άμεσων ξένων επενδύσεων κατά την εξέταση αντίστοιχων ομάδων χωρών προέλευσης και χωρών υποδοχής των επενδύσεων. Ο Akçay (2006) με χρήση cross sectional ανάλυσης που βασίστηκε σε δεδομένα από αναπτυσσόμενες χώρες απέτυχε να βρει οποιαδήποτε στατιστικά αποτελέσματα σχετικά με επίδραση της διαφθοράς στις εισροές άμεσων ξένων επενδύσεων και επιβεβαίωσε ότι, μεταξύ άλλων, η απελευθέρωση του εμπορίου είναι κρίσιμο καθοριστικό στοιχείο στη διαδικασία λήψης αποφάσεων των επενδυτών. Οι Wheeler και Mody (1992), αν και η εργασία τους δεν επικεντρώνεται στο θέμα της διαφθοράς όπως οι προηγούμενες, αλλά γενικότερα στις επιλογές περιοχών για επενδύσεις, δεν επισήμαναν επίσης κάποια αξία λόγου επίδραση της διαφθοράς στις ξένες επενδύσεις.

2.2.4 Μέθοδοι - Υποθέσεις

Μια κοινή μέθοδος με την οποία οι ερευνητές προσπάθησαν να μελετήσουν την παραπάνω σχέση είναι η εξέταση διμερών ροών άμεσων επενδύσεων και οι

διαφορές στα επίπεδα διαφθοράς ανάμεσα στις χώρες προέλευσης και προορισμού. Οι Bradaa, Drabek, Mendez, και Perez (2019), ανακάλυψαν πως οι εισροές άμεσων ξένων επενδύσεων επηρεάζονται αρνητικά από τη διαφθορά. Ένα σημαντικό πρόσθετο εύρημα στην εργασία τους είναι πως η διαφορά στα επίπεδα διαφθοράς ανάμεσα στις χώρες προέλευσης και προορισμού είναι ιδιαίτερα σημαντικός παράγοντας. Κάποιες εταιρίες αναπτύσσουν συγκεκριμένες τεχνογνωσίες με βάση το επιχειρηματικό περιβάλλον στις χώρες τους. Μετά αναζητούν χώρες με ανάλογες συνθήκες και επίπεδα διαφθοράς, γιατί εκεί θα αξιοποιήσουν καλύτερα την παραπάνω τεχνογνωσία τους, κάτι που δεν θα μπορούσαν να κάνουν στον ίδιο βαθμό σε χώρες με διαφορετικό περιβάλλον ή με τελείως διαφορετικές συνθήκες διαφθοράς. Οι Habib και Zurawicki (2002), σε μια παρόμοια μελέτη κατέληξαν στο ίδιο συμπέρασμα: Όσο μεγαλύτερη είναι η διαφορά στα επίπεδα διαφθοράς ανάμεσα στις χώρες προέλευσης και προορισμού, τόσο μικρότερες θα είναι οι εισροές άμεσων ξένων επενδύσεων. Στην εργασία τους ανέφεραν επίσης πως οι ξένοι επενδυτές συνήθως αποφεύγουν χώρες με πρόβλημα μεγάλης διαφθοράς, τόσο για ηθικούς όσο και για πρακτικούς λόγους. Όπως αναφέρουν οι Belgibayeva και Plekhanov (2015), οι διαφορές στα επίπεδα διαφθοράς ανάμεσα στις χώρες προέλευσης και προορισμού των επενδύσεων μπορεί να μεταβάλουν τον όγκο των εισροών άμεσων ξένων επενδύσεων σε περίπτωση που υπάρχουν μεταβολές στο θέμα της διαφθοράς στη χώρα προορισμού. Για παράδειγμα, υψηλά επίπεδα διαφθοράς μπορεί να μειώσουν τις εισροές επενδύσεων που προέρχονται από «καθαρότερες» χώρες, αλλά να αυξήσουν αυτές που προέρχονται από σχετικά διεφθαρμένες χώρες. Αν οι τελευταίες είναι μεγαλύτερες σε όγκο, τότε το αποτέλεσμα της διαφθοράς είναι θετικό για τη χώρα που τις υποδέχεται.

Είναι προφανές πως αυτό το αμφιλεγόμενο ζήτημα παραμένει ακόμα γόνιμο έδαφος για έρευνα. Πολλές εργασίες από τη σχετική βιβλιογραφία χρησιμοποιούν διαφορετικές μεθόδους και εργαλεία για την αποτίμηση της διαφθοράς και άλλων παραγόντων των μοντέλων τους, όπως η ποιότητα των θεσμών και η πολιτική σταθερότητα. Η παρούσα εργασία εξετάζει δεδομένα της περιόδου 2003 ως 2020 για 19 χώρες της ανατολικής Ευρώπης και 19 χώρες που ανήκουν στο υπόλοιπο της Ευρώπης. Οι περισσότερες εργασίες της βιβλιογραφίας στοχεύουν σε χώρες του τρίτου κόσμου χρησιμοποιώντας cross sectional models. Αυτό είναι πιθανόν να

οδηγήσει σε συμπεράσματα που πάσχουν από time-trends. Η περίοδος 18 ετών που διερευνάται στην παρούσα εργασία είναι αρκετή για να αποφευχθούν τέτοιου είδους προβλήματα στα επίπεδα διαφθοράς και στις εισροές άμεσων ξένων επενδύσεων παρέχοντας έτσι μια πιο ρεαλιστική εικόνα για κάθε χώρα. Η γνώση που αποκτήθηκε από την προηγούμενη σχετική βιβλιογραφία θα αξιοποιηθεί στη δόμηση των μεθόδων έρευνας στην παρούσα εργασία

Ειδικότερα προτείνω ένα μοντέλο που στοχεύει και απαντά σε συγκεκριμένους ελέγχους υποθέσεων, έχοντας ως στόχο την επιβεβαίωση ή μη των συμπερασμάτων της παραπάνω βιβλιογραφίας αλλά και στην εύρεση απαντήσεων στα θέματα που παραμένουν ανοιχτά με χρήση εναλλακτικών δεδομένων – proxies που βασίζονται στις επιδράσεις των βιομηχανιών.

Συνοψίζω παρακάτω τις εμπειρικές υποθέσεις που τα μοντέλα της εργασίας μου θα στοχεύσουν, με βάση τα ευρήματα και τις διαφωνίες στη σχετική βιβλιογραφία.

Υπόθεση 1 (Hypothesis 1, H1) - Οι εισροές άμεσων ξένων επενδύσεων θα αυξάνονται με την μείωση της διαφθοράς.

Υπόθεση 2 (Hypothesis 2, H2) - Μεγαλύτερο ποσοστό διεφθαρμένων βιομηχανιών οδηγεί σε μικρότερες εισροές άμεσων ξένων επενδύσεων.

3.Αναδρομή στη Θεωρία

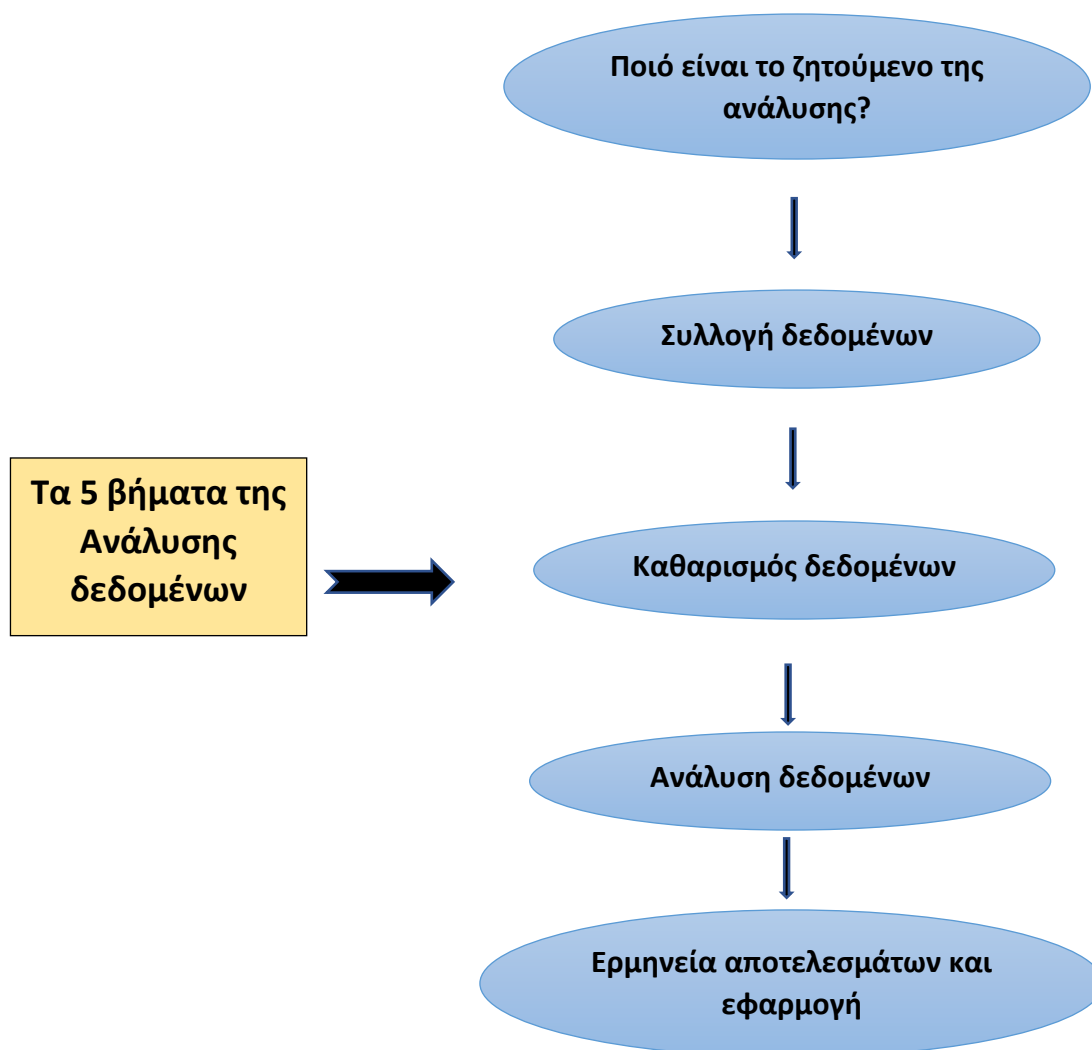
3.1 Ανάλυση Δεδομένων

Ανάλυση δεδομένων ονομάζεται η διαδικασία εύρεσης, συγκέντρωσης, καθαρισμού, εξέτασης και μοντελοποίησης δεδομένων, έτσι ώστε να εξαχθούν χρήσιμα συμπεράσματα και πληροφορίες με στόχο την καλύτερη λήψη αποφάσεων.

Αυτό που γίνεται γρήγορα κατανοητό από τον ορισμό είναι ότι η διαδικασία της ανάλυσης αποτελείται και η ίδια από διάφορα βήματα, τα οποία εάν ολοκληρωθούν με τον σωστό τρόπο φέρνουν το επιθυμητό αποτέλεσμα, δηλαδή αυτό της εξαγωγής συμπερασμάτων που θα λειτουργήσουν ως οδηγοί για μελλοντικές αποφάσεις. Η κατανόηση των βημάτων είναι αρκετά εύκολη, όμως αυτό που πρέπει να γίνει αντιληπτό είναι πως κάθε βήμα της ανάλυσης έχει την ίδια ακριβώς σημασία και δεν

πρέπει κανένα να υποτιμηθεί, καθώς αυτό ίσως να οδηγούσε σε μη ορθά συμπεράσματα και άρα λανθασμένες επιλογές.

Γράφημα 2: Τα 5 βήματα ανάλυσης δεδομένων



Βήμα 1^ο – Κατανόηση ζητούμενου ανάλυσης

Το πρώτο βήμα στην διαδικασία της ανάλυσης δεδομένων είναι η κατανόηση του σκοπού για τον οποίον αυτή θα γίνει. Αυτό είναι κάτι που πρέπει να ξεκαθαριστεί από την πλευρά ενός ερευνητή ή μίας επιχείρησης. Πιθανές ερωτήσεις που μπορεί να τεθούν είναι για παράδειγμα γιατί υπάρχει μείωση εσόδων ή/και μείωση του αριθμού πελατών σε μία περιοχή; Τι χρειάζεται για να αλλάξει αυτό; Τέτοιες ερωτήσεις που βοηθούν να καθοριστεί ο στόχος ή το ζητούμενο μπορεί να

ακούγονται αρχικά απλές, όμως πολλές φορές είναι αρκετά περίπλοκες και απαιτούν από τον αναλυτή των δεδομένων την πλήρη κατανόηση τους. Είναι ιδιαίτερα σημαντικό για τον αναλυτή σε περίπτωση που είναι μέλος μίας επιχείρησης ή ενός ομίλου να κατανοεί απόλυτα και τους στόχους αυτών, έτσι ώστε να μπορεί να θέσει το πρόβλημα σε σωστή βάση.

Βήμα 2^ο – Συλλογή Δεδομένων

Το επόμενο βήμα της αναλυτικής διαδικασίας είναι αυτό της συλλογής των δεδομένων. Αφού έχει καθοριστεί ο στόχος και ο σκοπός της ανάλυσης ξεκινάει η συγκέντρωση των ανάλογων δεδομένων. Σε αυτό το σημείο πρέπει να αποφασιστεί η στρατηγική της συλλογής, και μέρος αυτής είναι φυσικά το τι είδους δεδομένα είναι απαραίτητα για το εκάστοτε πρόβλημα. Τα δεδομένα μπορεί να είναι είτε ποιοτικά είτε ποσοτικά. Η φύση και η ποιότητα των δεδομένων που θα συλλεχθούν θα είναι εν τέλει και αυτή που θα κρίνει το πόσο εκτενής θα είναι η ανάλυση αλλά και το κατά πόσο είναι μαθηματικά ορθά τα αποτελέσματα και άρα τα τελικά συμπεράσματα.

Τα δεδομένα μπορούν να αντλούνται από τρεις κατηγορίες πηγών. Η πρώτη ονομάζεται κατηγορία πρώτου μέλους (first party) και αφορά δεδομένα από τον ίδιο τον οργανισμό ή την εταιρεία που πιθανόν να έχουν συλλεχθεί από τους ίδιους τους πελάτες της, όπως για παράδειγμα με μορφή ερωτηματολογίων, συνεντεύξεων ή φυσικά και με ιστορικά συναλλαγών τους. Αυτή η κατηγορία περιέχει κατά κύριο λόγο δεδομένα τα οποία είναι δομημένα και ευθέως καθορισμένα.

Η δεύτερη κατηγορία είναι αυτή του second party και αφορά συνήθως πηγές όπως το διαδίκτυο, τα κοινωνικά δίκτυα κ.α. Γίνεται κατανοητό πως αυτού του είδους οι πηγές βρίσκονται εκτός του οργανισμού ή της επιχείρησης, τα δεδομένα τους είναι τις περισσότερες φορές δομημένα, σχετικά αξιόπιστα και μπορούν λειτουργήσουν ως εναλλακτική πηγή πληροφορίας, έτσι ώστε να γίνει η ανάλυση πιο εκτενής και ακριβής από ότι θα ήταν με μία μόνο first party πηγή.

Η τρίτη κατηγορία ονομάζεται third party και αφορά τρίτους οργανισμούς που είναι σε θέση να παρέχουν πληροφορίες. Αυτές μπορεί να παρέχονται μέσω κυβερνητικών

πλατφορμών είτε από τον ιδιωτικό τομέα. Τέτοια παραδείγματα αποτελούν εταιρείες-όμιλοι με συμβουλευτικό χαρακτήρα, οι οποίοι παρέχουν τέτοιες υπηρεσίες έναντι χρηματικής αμοιβής. Σε αυτή τη κατηγορία συναντάμε αδόμητα δεδομένα πολύ συχνότερα από ότι στις άλλες δύο περιπτώσεις, αντίστοιχα με αυτά που συναντάμε στα μεγάλα δεδομένα (big data).

Βήμα 3ο – Καθαρισμός Δεδομένων

Το επόμενο βήμα μετά την συλλογή είναι ο καθαρισμός και η “τακτοποίηση” των δεδομένων. Αυτή είναι μία πολύ σημαντική διαδικασία καθώς δεν είναι όλα τα δεδομένα της συλλογής χρήσιμα για την μετέπειτα ανάλυση, έτσι η υπεύθυνη ομάδα πρέπει να δημιουργήσει ένα σύνολο μόνο με δεδομένα υψηλής ποιότητας και κατάλληλα για ανάλυση υψηλού επιπέδου. Χαρακτηριστικό της σημασίας αυτού του βήματος είναι το γεγονός πως οι αναλυτές ξοδεύουν τουλάχιστον το μισό χρόνο της αναλυτικής διαδικασίας στο στάδιο του καθαρισμού. Αυτός αποτελείται από διάφορες διεργασίες. Μία σειρά διεργασιών είναι ο καθαρισμός από διπλότυπες τιμές, outliers και λάθη. Αλλή μία σειρά είναι το δοθεί συγκεκριμένη δομή στο dataset όπως για παράδειγμα να διορθωθούν τυπογραφικά, οι υποδιαστολές να δηλώνονται με το ίδιο σύμβολο σε όλες τις στήλες κ.α. Βασικό στον καθαρισμό των δεδομένων είναι φυσικά να καθοριστεί ποια από τα δεδομένα έχουν πραγματική ουσία στην ανάλυση και ποια είναι απλά περιττά, χωρίς συνεισφορά στην ανάλυση. Τέλος, μέρος του καθαρισμού θεωρείται και η εύρεση ελλειπουσών τιμών και φυσικά ο κατάλληλος τρόπος για την κάλυψη τους.

Βήμα 4ο – Ανάλυση Δεδομένων

Μετά τον καθαρισμό, έρχεται η σειρά της ανάλυσης των δεδομένων. Σε αυτό το σημείο ο αναλυτής πρέπει να αποφασίσει για τον τύπο της ανάλυσης που πρέπει να πραγματοποιήσει και αυτός φυσικά εξαρτάται κάθε φορά από τον στόχο εκείνου για τον οποίο πραγματοποιείται η ανάλυση. Υπάρχουν διάφορα είδη ανάλυσης, όπως για παράδειγμα η ανάλυση χρονοσειρών, η ανάλυση παλινδρόμησης ή η ανάλυση συσταδοποίησης, γενικότερα όμως όλοι οι τύποι ανάλυσης ανήκουν σε κάποιες κατηγορίες, οι οποίες αναλύονται παρακάτω.

Διαγνωστική ανάλυση: επικεντρώνεται στο γιατί συνέβη κάτι, δηλαδή όπως λέει και το όνομα της δηλώνει την διάγνωση ενός προβλήματος. Ένα παράδειγμα μπορεί να είναι το γιατί οι πελάτες δεν προτίμησαν ένα νέο προϊόν έναντι του παλιού.

Περιγραφική ανάλυση: επικεντρώνεται στο τι συνέβη ήδη και συνήθως είναι το πρώτο βήμα πριν το γιατί συνέβη αυτό, δηλαδή λίγο πριν τη διαγνωστική ανάλυση. Δεν δίνει απάντηση δηλαδή στο γιατί αλλά θέτει τη βάση για περαιτέρω έρευνα, είναι δηλαδή το στάδιο που η εταιρεία πληροφορείται ότι οι πελάτες προτίμησαν το παλιό προϊόν και όχι το νέο.

Προγνωστική ανάλυση: αυτός ο τύπος ανάλυσης έχει ως στόχο να αναγνωρίζει μελλοντικές τάσεις έχοντας ως βάση ιστορικά δεδομένα. Αυτός ο τύπος ανάλυσης είναι πολύ σημαντικός για τομείς όπως ο χρηματοοικονομικός, αλλά και γενικότερα στον χώρο των επιχειρήσεων, αφού προβλέπει τάσεις, μελλοντικά αποτελέσματα κ.α. Με τη βοήθεια αλγορίθμων μηχανικής μάθησης, ο τύπος αυτός έχει αποκτήσει τεράστια σημασία καθώς προσφέρει εξαιρετικά ακριβή αποτελέσματα, εάν η ανάλυση πραγματοποιηθεί ορθά. Ο χρηματοοικονομικός και ασφαλιστικός τομέας βασίζονται σε τέτοιου είδους αναλύσεις.

Κανονιστική ανάλυση: στοχεύει στο να δίνει προτάσεις για το μέλλον, θεωρείται το τελικό στάδιο στην αναλυτική διαδικασία αλλά και το πιο περίπλοκο καθώς περιέχει πολλά στοιχεία και από τα άλλα είδη ανάλυσης που αναφέρθηκαν.

Βήμα 5ο – Ερμηνεία αποτελεσμάτων και συμπεράσματα:

Το τελευταίο βήμα της ανάλυσης δεδομένων είναι η ερμηνεία των αποτελεσμάτων που προέκυψαν στο προηγούμενο στάδιο. Αυτό είναι το στάδιο στο οποίο η κάθε εταιρεία συγκεντρώνει τη προσοχή της καθώς είναι το βήμα πριν τη λήψη των μέτρων και το στάδιο στο οποίο οι αναλυτές συνεργάζονται στενά με την κάθε εταιρεία. Μετά τα συμπεράσματα ακολουθούν δηλαδή οι προτάσεις για πιθανές λύσεις. Κατά τη διάρκεια της εξαγωγής των συμπερασμάτων αλλά και της επικοινωνίας των αποτελεσμάτων στους ενδιαφερόμενους είναι πολύ σημαντικό οι αναλυτές να εξηγούν όλα τα αποτελέσματα, θετικά ή αρνητικά, δηλαδή να μην διαλέγουν για

παράδειγμα μόνο τα “θετικά νέα”. Επίσης είναι κρίσιμο να εξηγούν τους περιορισμούς και τα προβλήματα που προέκυψαν κατά τη διάρκεια της ανάλυσης καθώς όλα μπορούν να παίξουν ρόλο στην λήψη των αποφάσεων.

3.2 Ανάλυση Παλινδρόμησης

Μεγάλο μέρος της παρούσας διπλωματικής θα βασιστεί σε μοντέλα ανάλυσης παλινδρόμησης. Τι είναι όμως ένα μοντέλο παλινδρόμησης; Η ανάλυση παλινδρόμησης είναι ένα από τα πιο σημαντικά εργαλεία που έχει να προσφέρει η επιστήμη της Στατιστικής. Η παλινδρόμηση λοιπόν με απλούς όρους είναι η περιγραφή και η αξιολόγηση της σχέσης μεταξύ μίας μεταβλητής και μίας άλλης ή άλλων (μεταβλητών). Με άλλα λόγια, η παλινδρόμηση στοχεύει στο να εξηγήσει “κινήσεις” σε μία μεταβλητή σε σχέση με κινήσεις άλλης ή άλλων μεταβλητών. Ορίζουμε ως y τη μεταβλητή της οποίας τις κινήσεις η παλινδρόμηση στοχεύει να εξηγήσει και τις μεταβλητές που χρησιμοποιούνται για να εξηγήσουν τη μεταβολή ως x_1, x_2, \dots, x_k . Μεταβολές δηλαδή σε k μεταβλητές οδηγούν σε αλλαγές σε κάποια άλλη μεταβλητή y . Φυσικά, οι ονομασίες των μεταβλητών ποικίλουν και τα x, y εδώ χρησιμοποιούνται ενδεικτικά.

3.2.1 Παλινδρόμηση και Συσχέτιση

Η συσχέτιση μεταξύ δύο μεταβλητών μετρά τον βαθμό γραμμικής σχέσης μεταξύ τους. Έαν υπάρχει συσχέτιση μεταξύ μίας μεταβλητής x και μίας μεταβλητής y , σημαίνει ότι η x και η y συμπεριφέρονται με ένα εντελώς συμμετρικό τρόπο. Έτσι, αυτό δε σημαίνει ότι μία αλλαγή στην x θα φέρει αλλαγή και στην y . Αντίθετα δείχνει ότι υπάρχει μεταξύ των μεταβλητών γραμμική σχέση και ότι οι κινήσεις τους είναι συσχετιζόμενες στο βαθμό που δίνεται από τον συντελεστή συσχέτισης.

Στη παλινδρόμηση, η εξαρτημένη μεταβλητή (y) και οι ανεξάρτητες έχουν μία τελείως διαφορετική σύνδεση μεταξύ τους. Η y θεωρείται ότι είναι τυχαία ή στοχαστική κατά κάποιο τρόπο, δηλαδή έχει κατανομή πιθανότητας. Οι μεταβλητές x θεωρούνται ότι παίρνουν μη στοχαστικές τιμές σε επαναλαμβανόμενα δείγματα. Η παλινδρόμηση θεωρείται πιο ευέλικτη σαν εργαλείο αλλά και πιο ισχυρή από την συσχέτιση.

3.2.2 Απλή Γραμμική Παλινδρόμηση

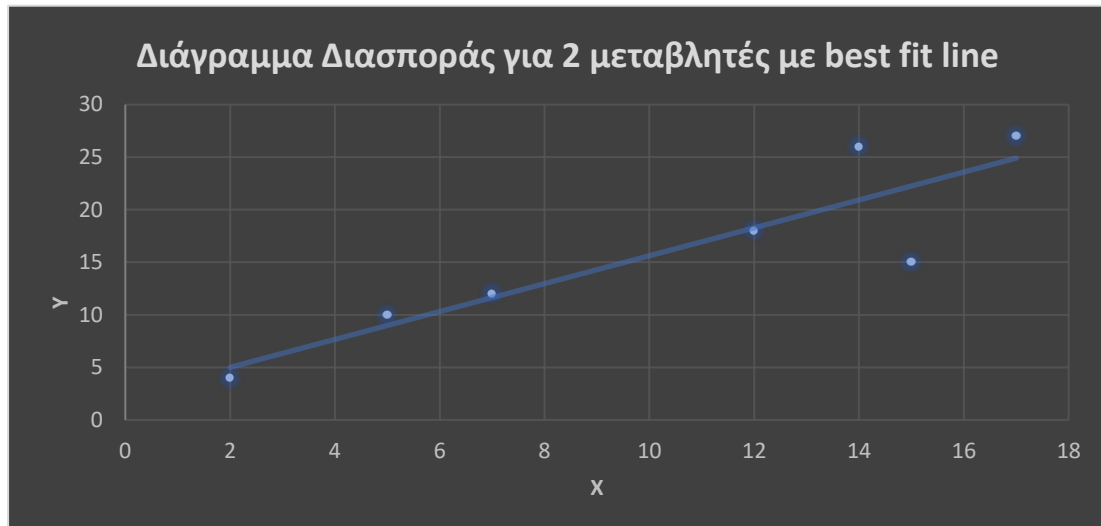
Η πιο απλή μορφή παλινδρόμησης είναι αυτή της απλής γραμμικής παλινδρόμησης. Σε αυτή υπάρχει φυσικά μία εξαρτημένη μεταβλητή αλλά και μόνο μία ανεξάρτητη μεταβλητή. Η πρώτη μπορεί να προσεγγιστεί με μία γραμμική συνάρτηση του X . Πιο συγκεκριμένα εάν ένας ερευνητής θέλει να βγάλει κάποια συμπεράσματα για τη σχέση μεταξύ δύο μεταβλητών, όπως για παράδειγμα το εάν μία αύξηση της μεταβλητής x φέρνει αύξηση στην μεταβλητή y , το πρώτο πράγμα που θα έκανε θα ήταν η κατασκευή ενός διαγράμματος διασποράς. Εάν υποθέσουμε ότι το διάγραμμα διασποράς δείχνει προσεγγιστικά πάντα ότι υπάρχει γραμμική σχέση μεταξύ των μεταβλητών x και y , για παράδειγμα θετική σχέση, όπου η αύξηση στη x φέρνει αύξηση στη y , τότε αυτή η σχέση μπορεί να περιγραφεί επίσης προσεγγιστικά από μία ευθεία γραμμή. Φυσικά ένας τρόπος θα ήταν να σχεδιαστεί με το χέρι μία ευθεία που να ταιριάζει με τα δεδομένα. Έτσι αντίστοιχα και προσεγγιστικά θα βρίσκαμε την κλίση και την ευθεία. Όμως αυτός ο τρόπος μπορεί να δημιουργήσει αποτελέσματα τα οποία είναι αρκετά ανακριβή. Εξαιτίας αυτού, θα ήταν ιδιαίτερα χρήσιμο αν η σχέση μεταξύ των μεταβλητών μπορούσε να περιγραφεί από μία εξίσωση που μπορεί να υπολογιστεί με μία συγκεκριμένη διαδικασία. Έτσι για την εύρεση της ευθείας γραμμής που προκύπτει με βάση τα δεδομένα χρησιμοποιείται η παρακάτω εξίσωση:

$$y = a + \beta x$$

Το επόμενο βήμα είναι ο ερευνητής να βρει τις κατάλληλες τιμές για τις παραμέτρους a και β , έτσι ώστε η γραμμή να βρίσκεται όσο πιο κοντά γίνεται στα δεδομένα συνολικά. Βέβαια αυτό το μοντέλο δεν είναι απόλυτα ρεαλιστικό, καθώς υποθέτει πως εάν οι τιμές a και β υπολογιστούν, τότε με δεδομένη την τιμή του x , θα μπορούσε με απόλυτη βεβαιότητα να υπολογιστεί και η ακριβής τιμή του y . Αυτό αυτονόητα απέχει από την πραγματικότητα. Στον χώρο της στατιστικής αυτό θα σημαίνει ότι το μοντέλο ταιριάζει στα δεδομένα απόλυτα, δηλαδή ότι όλα τα δεδομένα παρατηρήσεις βρίσκονται πάνω ακριβώς στην ευθεία γραμμή. Για να γίνει το μοντέλο πιο ρεαλιστικό θα πρέπει να ενσωματωθεί σε αυτό ένας τυχαίος όρος που θα συμβολίζει το σφάλμα, που εδώ θα συμβολιστεί με u :

$$y_t = \alpha + \beta x_t + u_t$$

Στην νέα εξίσωση το t συμβολίζει τον αριθμό των παρατηρήσεων. Στον όρο του σφάλματος ενσωματώνονται χαρακτηριστικά όπως για παράδειγμα τα λάθη που μπορεί να υπάρχουν στον υπολογισμό του y και τα οποία δεν μπορούν να



μοντελοποιηθούν. Άλλα χαρακτηριστικά είναι οι τυχαίες εξωτερικές επιδράσεις, οι οποίες επίσης δε μπορούν να μοντελοποιηθούν. Είναι φανερό πως η σημασία αυτού του όρου είναι μεγάλη και πολύπλευρη. Σε αυτό το στάδιο τίθεται το ερώτημα του πως αποφασίζεται το ποιες είναι οι κατάλληλες τιμές για τις παραμέτρους α και β . Το α και το β επιλέγονται έτσι ώστε οι κάθετες αποστάσεις από τα σημεία των δεδομένων προς την ευθεία να ελαχιστοποιούνται, έτσι ώστε η γραμμή να προσεγγίζει τα δεδομένα όσο καλύτερα γίνεται. Έτσι λοιπόν οι παράμετροι επιλέγονται με αυτό το στόχο, δηλαδή την συνολική ελαχιστοποίηση των καθέτων αποστάσεων.

Η επιλογή λοιπόν του να τραβήξει κανείς μία γραμμή προσεγγιστικά, αρκεί μόνο κάποιες φορές όταν υπάρχει ανάγκη μόνο για ορισμένα ενδεικτικά αποτελέσματα και είναι αυτονόητο ότι δεν μπορεί κανένας ερευνητής να στηριχθεί σε αυτά καθώς πιθανότατα το μοντέλο δεν θα είναι ακριβές. Η πιο δημοφιλής και συνηθισμένη μέθοδος για να προσαρμόσει κανείς την ευθεία στα δεδομένα, που δεν βασίζεται στην οπτική αλλά στην στατιστική και στα οικονομετρικά μοντέλα, είναι η μέθοδος ελαχίστων τετραγώνων (OLS). Η μέθοδος OLS βασίζεται στις κάθετες αποστάσεις από

το κάθε σημείο μέχρι τη γραμμή, στην συνέχεια στην ύψωση τους στο τετράγωνο και τέλος στην ελαχιστοποίηση του συνόλου του εμβαδού των τετραγώνων.

Το πρώτο βήμα είναι να συμβολιστεί το πραγματικό σημείο της παρατήρησης y με y_t και \hat{y}_t να συμβολιστεί η προσαρμοσμένη τιμή απο την γραμμή παλινδρόμησης. Με άλλα λόγια δεδομένης μίας τιμής x της παρατήρησης t , \hat{y}_t είναι η τιμή που θα έχει προβλέψει το μοντέλο για την y . Τέλος, συμβολίζεται με \hat{u}_t το υπόλοιπο-σφάλμα, δηλαδή η διαφορά ανάμεσα στη πραγματική τιμή της y και την προσαρμοσμένη τιμή που βρέθηκε από το μοντέλο για το συγκεκριμένο σημείο, δηλαδή $\hat{u}_t = (y_t - \hat{y}_t)$. Ο στόχος δηλαδή σε αυτή την περίπτωση είναι η ελαχιστοποίηση του συνόλου των τετραγώνων των υπολοίπων (\hat{u}_t^2).

Η ελαχιστοποίηση του συνόλου των τετραγωνικών αποστάσεων για τη περίπτωση τεσσάρων παρατηρήσεων δίνεται από την ελαχιστοποίηση του:

$$(\hat{u}_1^2 + \hat{u}_2^2 + \hat{u}_3^2 + \hat{u}_4^2)$$

Το σύνολο αυτό ονομάζεται ως άθροισμα τετραγώνων των σφαλμάτων - RSS.

Αν $\hat{\alpha}$ και $\hat{\beta}$ συμβολίζουν τις τιμές του α και β που επελέγησαν απο την ελαχιστοποίηση του RSS, τότε η εξίσωση για την προσαρμοσμένη ευθεία είναι η ακόλουθη: $\hat{y}_t = \hat{\alpha} + \hat{\beta}x_t$.

Στην συνέχεια, αν συμβολιστεί με L το RSS, που ονομάζεται αλλιώς και συνάρτηση κόστους και έχουμε T παρατηρήσεις προκύπτει η παρακάτω σχέση:

$$L = \sum_{t=1}^T (y_t - \hat{y}_t)^2 = \sum_{t=1}^T (y_t - \hat{\alpha} + \hat{\beta}x_t)^2$$

Το L ελαχιστοποιείται ως προς το $\hat{\alpha}$ και το $\hat{\beta}$ για να βρεθούν οι τιμές του α και β οι οποίες με τη σειρά τους ελαχιστοποιούν το RSS έχοντας ως τελικό στόχο να βρεθεί η ευθεία που είναι πλησιέστερα στα δεδομένα. Το πρώτο βήμα είναι να θέσουμε τις πρώτες παραγώγους ίσες με μηδέν. Οι εκτιμητές για την ευθεία και την κλίση δίνονται απο τις παρακάτω σχέσεις.

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

$$\hat{\beta} = \frac{\sum(x_t - \bar{x})(y_t - \bar{y})}{\sum(x_t - \bar{x})^2}$$

Όπου σε αυτή τη μορφή το $\hat{\beta}$ ισούται με τη δειγματική συνδιακύμανση μεταξύ x και y διαιρούμενη από τη δειγματική διακύμανση της x . Αυτός ο τρόπος λοιπόν ονομάζεται μέθοδος ελαχίστων τετραγώνων (OLS) και βρίσκει τις τιμές των παραμέτρων που ταιριάζουν καλύτερα στα δεδομένα. Οι υπόλοιπες τιμές για τις παραμέτρους θεωρείται ότι θα οδηγούσαν σε χειρότερη προσαρμογή του μοντέλου, δηλαδή σε υψηλότερο RSS.

3.2.3 Υποθέσεις απλού γραμμικού μοντέλου

Το απλό γραμμικό μοντέλο παλινδρόμησης βασίζεται σε πέντε βασικές υποθέσεις. Τα δεδομένα για το x_t είναι παρατηρήσιμα, αλλά επειδή το y_t εξαρτάται από το u_t είναι απαραίτητο να ξεκαθαριστεί το πως προέκυψε το u_t . Για αυτό το λόγο οι συγκεκριμένες υποθέσεις αφορούν κυρίως τον όρο u_t , δηλαδή το μη παρατηρήσιμο λάθος. Ακολουθούν οι πέντε υποθέσεις:

1. Τα σφάλματα έχουν μέση τιμή 0 $\Rightarrow E(u_t) = 0$
2. Η διακύμανση των σφαλμάτων είναι σταθερή και πεπερασμένη για όλες τις τιμές του $x_t \Rightarrow \text{var}(u_t) = \sigma^2 < \infty$
3. Τα σφάλματα είναι μεταξύ τους γραμμικώς ανεξάρτητα $\Rightarrow \text{cov}(u_i, u_j) = 0$
4. Όχι σχέση μεταξύ του σφάλματος και του αντίστοιχου $x \Rightarrow \text{cov}(u_t, x_t) = 0$
5. Το u_t είναι κανονικά κατανομημένο $\Rightarrow u_t = N(0, \sigma^2)$

3.2.4 Έλεγχος υποθέσεων απλού γραμμικού μοντέλου

Στην ερμηνεία που δόθηκε στην παράμετρο β στις προηγούμενες υποενότητες και κατά τη διάρκεια της περιγραφής του απλού γραμμικού μοντέλου, είδαμε πως η παράμετρος αυτή συμβολίζει την αναμενόμενη μεταβολή στην μεταβλητή Y , όταν η ανεξάρτητη μεταβλητή μεταβληθεί κατά μία μονάδα. Αν υποθέσουμε ότι $\beta_1 = 0$, αυτό σημαίνει ότι η μέση τιμή της Y ($E(Y)$) δεν μεταβάλλεται όταν αλλάζει η τιμή της X .

Είναι δηλαδή μία περίπτωση που σημαίνει ότι η τιμή που θα λάβει η X κάθε φορά δεν μας πληροφορεί για κάτι σχετικά με την αναμενόμενη τιμή της Y .

Αν γράψουμε το απλό γραμμικό μοντέλο στην παρακάτω μορφή

$$Y = \beta_0 + \beta_1 x + \varepsilon$$

το ε είναι τυχαία μεταβλητή με μέση τιμή ίση με 0 και διακύμανση σ^2 . Με $\beta_1 = 0$ καταλήγουμε πως $Y = \beta_0 + \varepsilon$. Εδώ η τυχαία μεταβλητή Y παίρνει τιμές γύρω από τη μέση τιμή της $E(Y) = \beta_0$ και δεν εξαρτάται από την τιμή που παίρνει η μεταβλητή X . Καταλήγοντας, είναι εμφανές πως για να δούμε αν η X μπορεί να αξιοποιηθεί για την πρόβλεψη της Y , αυτό που πρέπει να αποφασιστεί είναι το αν δεχόμαστε πως $\beta_1 = 0$ ή όχι. Για να αποφασιστεί αν ισχύει αυτή η στατιστική υπόθεση θα εφαρμοστούν έλεγχοι υποθέσεων. Η παρούσα εργασία στηρίζεται σε τέτοιες υποθέσεις και για αυτό η αναφορά πάνω στη θεωρία τέτοιων ελέγχων είναι απαραίτητη. Ο έλεγχος βασισμένη στη περίπτωση που αναφέρθηκε προηγουμένως γράφεται ως εξής:

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 \neq 0$$

Η πρώτη υπόθεση, δηλαδή η H_0 , ονομάζεται μηδενική και η H_1 εναλλακτική. Στη περίπτωση του απλού γραμμικού μοντέλου αν γίνει αποδεκτή η μηδενική υπόθεση τότε αυτό σημαίνει πως η ανεξάρτητη μεταβλητή X δεν μπορεί να συνεισφέρει στην πρόβλεψη της τιμής της εξαρτημένης μεταβλητής Y . Αντιθέτως αν απορριφθεί η μηδενική υπόθεση τότε αυτό σημαίνει πως η X είναι στατιστικώς σημαντική για την πρόβλεψη της τιμής της Y . Η απόφαση αυτή λαμβάνεται με βάση τις τιμές x_1, x_2, \dots, x_n που δίνονται στην εκάστοτε ανεξάρτητη μεταβλητή X και τις τιμές που θα προκύψουν για τις τυχαίες μεταβλητές Y_1, Y_2, \dots, Y_n . Για να σχηματιστεί ο κανόνας της απόφασης πρέπει να οριστεί η κατανομή των Y_i , $i = 1, 2, \dots, n$, Δεδομένης μίας παραμέτρου θ κατασκευάζουμε διάστημα εμπιστοσύνης με συντελεστή $1-\alpha$ και στόχος είναι να ελεγχθεί η υπόθεση $H_0 : \theta = c$ ή η εναλλακτική υπόθεση $H_1 : \theta \neq c$. Η απόφαση θα βασιστεί στον παρακάτω κανόνα:

Σε περίπτωση που η τιμή c δεν βρίσκεται μέσα στο διάστημα εμπιστοσύνης για τη παράμετρο θ με συντελεστή εμπιστοσύνης $1-\alpha$ τότε η μηδενική υπόθεση θα

απορριφθεί έναντι της εναλλακτικής, πάντα σε επίπεδο σημαντικότητας α . Αντίθετα αν ανήκει στο διάστημα δεν απορρίπτεται η μηδενική υπόθεση. Καταλαβαίνουμε δηλαδή ότι το βασικότερο σε αυτή τη περίπτωση είναι να ελεγχθεί το πόσο η τιμή $c = 0$ βρίσκεται εντός ή εκτός του διαστήματος. Πρέπει να ελεγχθεί δηλαδή το αν ισχύουν τα παρακάτω:

$$0 < \widehat{\beta}_1 - s(\widehat{\beta}_1) t_{v-2}(\alpha/2) \quad \text{ή} \quad \widehat{\beta}_1 + s(\widehat{\beta}_1) t_{v-2}(\alpha/2) < 0$$

Από τα οποία καταλήγουμε στο εξής

$$\frac{\widehat{\beta}_1}{s(\widehat{\beta}_1)} > t_{v-2}(\alpha/2) \quad \text{ή} \quad \frac{\widehat{\beta}_1}{s(\widehat{\beta}_1)} < -t_{v-2}(\alpha/2)$$

Έτσι καταλήγουμε στον παρακάτω συμβολισμό

$$T = \frac{\widehat{\beta}_1}{s(\widehat{\beta}_1)}$$

Από τα παραπάνω μπορούμε τώρα για τον κανόνα απόφασης σε ένα επίπεδο σημαντικότητας α να ορίσουμε πως αν ισχύει ότι $|T| > t_{v-2}(\alpha/2)$, η μηδενική υπόθεση θα απορριφθεί, ενώ αν ισχύει πως $|T| \leq t_{v-2}(\alpha/2)$, τότε η μηδενική υπόθεση δεν απορρίπτεται. Η περιοχή της ανισότητας $|T| > t_{v-2}(\alpha/2)$ ονομάζεται περιοχή απόρριψης.

Αντίστοιχα, υπάρχει φυσικά και η πιθανότητα να δημιουργηθούν και λάθος συμπεράσματα ως προς την επιλογή απόρριψης της μηδενικής υπόθεσης. Η πιθανότητα αυτή δίνεται από τον παρακάτω τύπο

$$P[|T| > t_{v-2}(\alpha/2) | H_0] = P[|T| > t_{v-2}(\alpha/2) | \beta_1 = 0]$$

Αλλά αφού

$$\begin{aligned} P[|T| > t_{v-2}(\alpha/2) | \beta_1 = 0] &= 1 - P[|T| \leq t_{v-2}(\alpha/2) | \beta_1 = 0] = \\ &= 1 - P[-t_{v-2}(\alpha/2) \leq \frac{\widehat{\beta}_1}{s(\widehat{\beta}_1)} \leq t_{v-2}(\alpha/2) | \beta_1 = 0] \end{aligned}$$

Άρα καταλήγουμε πως $P[|T| > t_{v-2}(\alpha/2) | H_0] = 1 - (1 - \alpha) = \alpha$

Όπου α ονομάζεται το επίπεδο σημαντικότητας και ισούται με την πιθανότητα λάθος απόρριψης της H_0 .

Η ίδια διαδικασία μπορεί να ακολουθηθεί και για την κατασκευή του κανόνα απόφασης για τον παρακάτω έλεγχο:

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 \neq 0$$

Με το T αυτή τη φορά να δίνεται απο την παρακάτω σχέση:

$$T = \frac{\widehat{\beta}_0}{s(\widehat{\beta}_0)}$$

Τέτοιου είδους έλεγχοι ονομάζονται αμφίπλευροι.

Υπάρχουν όμως περιπτώσεις-προβλήματα που ένας έλεγχος αυτής της μορφής μπορεί να μην εξυπηρετεί. Υπάρχουν δηλαδή προβλήματα που οι παράμετροι β_0 και β_1 δεν μπορούν να πάρουν αρνητικές τιμές. Σε αυτή τη περίπτωση ο έλεγχος θα ήταν της παρακάτω μορφής

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 > 0$$

Και

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 > 0$$

Αντίστοιχα υπάρχουν περιπτώσεις όπου β_0 και β_1 δεν λαμβάνουν θετικές τιμές και για αυτές τις περιπτώσεις ο κανονικός έλεγχος υποθέσεων παίρνει τη εξής μορφή

$$H_0 : \beta_1 = 0$$

$$H_1 : \beta_1 < 0$$

Και

$$H_0 : \beta_0 = 0$$

$$H_1 : \beta_0 < 0$$

Οι δύο τελευταίες περιπτώσεις δεν ανήκουν στη κατηγορία των αμφίπλευρων ελέγχων αλλά ονομάζονται μονόπλευροι έλεγχοι. Η απόφαση για αυτούς τους ελέγχους είναι αρκετά παρόμοια με αυτήν που παρουσιάστηκε ήδη, μόνο που σε αυτήν τη περίπτωση δε χρησιμοποιείται η απόλυτη τιμή ενώ τροποποιείται και το σημείο κατανομής. Για την ακρίβεια ο έλεγχος για ένα επίπεδο σημαντικότητας α βασίζεται στο ότι σε περίπτωση που ισχύει ότι $T > t_{\nu-2}(\alpha)$ η μηδενική υπόθεση απορρίπτεται, ενώ στην περίπτωση όπου $T \leq t_{\nu-2}(\alpha)$, η H_0 δεν απορρίπτεται. Αντίστοιχα για τη δεύτερη περίπτωση απορρίπτουμε την μηδενική υπόθεση εάν $T < -t_{\nu-2}(\alpha)$ και δεν την απορρίπτουμε εάν $T \geq -t_{\nu-2}(\alpha)$

3.3 Τύποι Δεδομένων

Υπάρχουν γενικά τρεις κατηγορίες δεδομένων που μπορούν να χρησιμοποιηθούν στην ποσοτική ανάλυση των οικονομικών προβλημάτων: Δεδομένα Time-series, cross-sectional δεδομένα και panel δεδομένα. Κάθε μια από αυτές τις κατηγορίες θα συζητηθεί με τη σειρά της στη συνέχεια, αλλά πρώτα αξίζει να αναφερθεί ένα άλλο χαρακτηριστικό των δεδομένων που είναι σημαντικό, ο βαθμός ομαδοποίησης - degree of aggregation τους. Πολλές μορφές δεδομένων ξεκινούν σαν μεμονωμένες παρατηρήσεις αλλά στη συνέχεια ομαδοποιούνται. Για παράδειγμα, μπορούμε να καταγράψουμε την τιμή πώλησης ενός συγκεκριμένου ακινήτου σε έναν συγκεκριμένο δρόμο κάθε φορά που πωλείται και να παρατηρήσουμε πόσο και γιατί μεταβάλλεται διαχρονικά. Αλλά συνήθως, ένα συγκεκριμένο σπίτι είναι απίθανο να πουληθεί συχνότερα από μια φορά κάθε πέντε ή δέκα χρόνια. Έτσι, είναι συνηθισμένο να σχηματίζουμε δείκτες τιμών σπιτιών, που καταμετρούν τον μέσο όρο των τιμών των σπιτιών που πωλούνται μέσα σε μια συγκεκριμένη χρονική περίοδο, π.χ σε ένα μήνα. Έτσι οι μεμονωμένες τιμές πώλησης ενός αριθμού σπιτιών κατά κάποιο τρόπο συνδυάζονται (ομαδοποιούνται) και μετασχηματίζονται σε ένα δείκτη. Οι τιμές των σπιτιών μπορούν να ομαδοποιηθούν σε επίπεδο δρόμου, περιοχής, πόλης ή χώρας. Με τους δείκτες αυτούς μπορεί να ξεπεραστεί το πρόβλημα του ότι

τα μεμονωμένα σπίτια δεν πωλούνται τόσο συχνά, με τον συνδυασμό πληροφοριών από πωλήσεις πολλών διαφορετικών σπιτιών. Παρόλα αυτά, ο ερευνητής που δημιουργεί τους δείκτες αυτούς πρέπει να προσέχει ώστε να συγκρίνει όμοια αντικείμενα, και να προσαρμόζει τα δεδομένα με τέτοιο τρόπο που να λαμβάνονται υπόψη οι παραλλαγές στα είδη των σπιτιών που πωλούνται.

Είναι πιθανό να μας ενδιαφέρουν τα εθνικά δεδομένα, οπότε θα χρειαζόμασταν ένα εθνικό δείκτη. Για παράδειγμα, ίσως μας ενδιέφερε να γνωρίζουμε ποιοί παράγοντες προκαλούν πτώση των τιμών των σπιτιών στην Ελλάδα συνολικά σε μια συγκεκριμένη χρονική περίοδο, οπότε οι τιμές των μεμονωμένων σπιτιών ή οι τιμές σε κάποιες συγκεκριμένες πόλεις είναι αδιάφορες.

Σημειώνεται πως με τη χρήση ομαδοποιημένων δεδομένων έχουμε τη γενική εικόνα, αλλά χάνουμε πολλές λεπτομέρειες. Για παράδειγμα, μπορεί να βλέπουμε αύξηση των τιμών σε ολόκληρη την Ελλάδα αλλά αν μελετήσουμε τα επι μέρους στοιχεία για Αττική, Μακεδονία, Θεσσαλία, Πελοπόννησο, Κρήτη κλπ μπορεί να διαπιστώσουμε πως άνοδος τιμών υπάρχει μόνο στην Αττική, ενώ στις άλλες περιφέρειες οι τιμές πέφτουν. Αλλά επειδή ο όγκος των πωλήσεων σπιτιών στην Αττική είναι πολύ μεγαλύτερος, ο μέσος όρος εμφανίζεται ανοδικός.

3.3.1 Time-Series Data

Τα δεδομένα Time-series, όπως υποδηλώνει το όνομα, είναι δεδομένα που αφορούν μια χρονική περίοδο σε σχέση με μία ή περισσότερες μεταβλητές. Τα δεδομένα χρονοσειρών είναι συσχετισμένα με συγκεκριμένη συχνότητα παρατηρήσεων ή συχνότητα συγκέντρωσης data points. Η συχνότητα είναι απλά ένα μέτρο της κανονικότητας και του διαστήματος μέσα στο οποίο συλλέγονται ή καταγράφονται τα δεδομένα.

Πολλά οικονομικά στοιχεία δεν πρωτοεμφανίζονται κανονικά καταναμημένα στο χρόνο. Για παράδειγμα, η τιμή μιας κοινής μετοχής συγκεκριμένης εταιρίας μπορεί να καταγράφει αλλαγές όποτε πραγματοποιούνται αγοραπωλησίες της. Τέτοιες κινήσεις και καταγραφές είναι πολύ απίθανο να πραγματοποιούνται μέσα σε

σταθερά χρονικά διαστήματα, π.χ μπορεί να μην υπάρχει δραστηριότητα ανάμεσα στην ώρα που κλείνει ή αγορά και την ώρα που ξανανοίγει την επόμενη μέρα. Υπάρχει επίσης κατά κανόνα μικρότερη δραστηριότητα γύρω στις ώρες ανοίγματος και κλεισίματος, καθώς και τις ώρες μεσημεριανού φαγητού. Αν και υπάρχουν διάφοροι τρόποι να αντιμετωπιστεί το ζήτημα αυτό, μια συνηθισμένη και απλή προσέγγιση είναι να επιλεγεί μια κατάλληλη συχνότητα, και να χρησιμοποιηθεί για το συγκεκριμένο χρονικό διάστημα η τελευταία επικρατούσα τιμή.

Είναι επίσης γενικά απαιτούμενο όλα τα δεδομένα ενός μοντέλου να προέρχονται από την ίδια συχνότητα παρατήρησης. Έτσι, για παράδειγμα, παλινδρομήσεις που επιχειρούν να υπολογίσουν ένα μοντέλο τιμολόγησης για arbitrage με χρήση μηνιαίων παρατηρήσεων μακροοικονομικών παραγόντων, πρέπει επίσης να χρησιμοποιήσουν και μηνιαία στοιχεία για την απόδοση μετοχών, ακόμα και αν υπάρχουν διαθέσιμα ημερήσια ή εβδομαδιαία.

Τα δεδομένα μπορεί να είναι ποσοτικά (π.χ ισοτιμίες, τιμές, αριθμός μετοχών σε κυκλοφορία) ή ποιοτικά (π.χ η ημέρα της εβδομάδας, μια έρευνα για τα οικονομικά προϊόντα που αγοράστηκαν σε μια χρονική περίοδο, κάποια αξιολόγηση φερεγγυότητας, κλπ).

Παραδείγματα προβλημάτων που μπορούν να αντιμετωπιστούν με χρήση time series data είναι : το πώς το ύψος των χρηματιστηριακών δεικτών μιας χώρας σχετίζεται με τα βασικά μακροοικονομικά της μεγέθη ή το πως η τιμή της μετοχής μίας εταιρίας μεταβλήθηκε μετά την ανακοίνωση της πληρωμής μερισμάτων.

Σε όλες τις παραπάνω περιπτώσεις είναι σαφές πως η διάσταση του χρόνου είναι η σημαντικότερη, και η ανάλυση θα γίνει με χρήση των τιμών των μεταβλητών μέσα στο χρόνο.

3.3.2 Cross-Sectional Data

Cross-Sectional είναι τα δεδομένα πάνω σε μια ή περισσότερες μεταβλητές που συγκεντρώνονται σε ένα μοναδικό χρονικό σημείο. Για παράδειγμα, τα δεδομένα

μπορεί να αφορούν, μια έρευνα για τη χρήση χρηματιστηριακών υπηρεσιών του Διαδικτύου, ένα δείγμα αποδόσεων του χρηματιστηρίου της Νέας Υόρκης και ένα δείγμα αξιολόγησης φερεγγυότητας των ομολόγων Βρετανικών τραπεζών. Προβλήματα που μπορούν να αντιμετωπιστούν με χρήση Cross-Sectional δεδομένων είναι η σχέση ανάμεσα στο μέγεθος μίας εταιρίας και στην απόδοση των μετοχών της ή ακόμα και η συσχέτιση του ΑΕΠ μίας χώρας με την πιθανότητα η χώρα να πτωχεύσει.

3.3.3 Panel Data

Τα Panel Data, γνωστά επίσης και σαν longitudinal, έχουν τις διαστάσεις τόσο των δεδομένων χρονοσειρών όσο και των συγχρονικών δεδομένων. Για παράδειγμα, οι ημερήσιες τιμές ενός αριθμού μετοχών blue chips σε διάστημα δύο ετών.

Ευτυχώς σχεδόν όλες οι τεχνικές και αναλύσεις της οικονομετρίας είναι το ίδιο αξιόπιστες προκειμένου είτε για δεδομένα time-series είτε για cross-sectional. Για δεδομένα time-series συνηθίζεται να δηλώνουμε τα individual observation numbers με χρήση του δείκτη t και τον συνολικό αριθμό των παρατηρήσεων που είναι διαθέσιμες για ανάλυση με το T . Για cross-sectional δεδομένα τα individual observation numbers δηλώνονται με χρήση του δείκτη i και τον συνολικό αριθμό των παρατηρήσεων που είναι διαθέσιμες για ανάλυση με το N . Σημειώνεται πως, αντίθετα με την περίπτωση time-series, δεν υπάρχει φυσική διάταξη των παρατηρήσεων σε ένα cross-sectional δείγμα. Για παράδειγμα, οι παρατηρήσεις i μπορεί να αφορούν τιμές μετοχών διαφορετικών εταιριών σε συγκεκριμένο χρονικό σημείο, σε αλφαβητική σειρά με βάση τα ονόματά τους. Ετσι, στην περίπτωση των cross-sectional δεδομένων, είναι απίθανο να περιέχεται κάποια χρήσιμη πληροφορία στο γεγονός πως η Barclays ακολουθεί την Banco Santander σε ένα δείγμα αξιολογήσεων πιστοληπτικής ικανότητας, μια και είναι τελείως συμπτωματικό το ότι τα ονόματα και των δύο αρχίζουν με το ίδιο γράμμα. Αντίθετα, στο πλαίσιο των time-series η διάταξη των δεδομένων είναι σημαντική καθώς αυτά είναι συνήθως σε χρονική σειρά.

Τέλος, ένα είδος δεδομένων που είναι κάπως διαφορετικά από τα παραπάνω είναι τα pooled cross-section και time-series. Αυτό συμβαίνει όταν η μεταβλητή που μας ενδιαφέρει έχει διαστάσεις τόσο time-series όσο και cross-section, αλλά για κάποιο λόγο δεν χρησιμοποιούμε τα χαρακτηριστικά αυτά αλλά απλά συνδιάζουμε όλες τις παρατηρήσεις μαζί. Για παράδειγμα, μπορεί να έχουμε μηνιαία δεδομένα δέκα ετών για τα κέρδη που πέτυχαν έξι διαφορετικοί επενδυτές σε μια ομάδα, αλλά αν αγνοήσουμε τη χρονική σειρά των δεδομένων καθώς και το τι κέρδη πέτυχε ο κάθε μεμονωμένος επενδυτής και απλά βάλουμε όλα τα κέρδη σε μια αταξινόμητη στήλη, αυτό θα είναι ένα pooled δείγμα. Δεν θα είναι πίνακας δεδομένων, μια και δεν θα είμαστε σε θέση να παρατηρήσουμε την επίδοση του κάθε μεμονωμένου επενδυτή κατα μήνα. Στην πράξη, τα pooled δεδομένα αντιμετωπίζονται σαν να είναι απλά ένα μεγαλύτερο cross-sectional δείγμα.

3.3.4 Continuous και Discreet data - Συνεχή και Διακριτά

Παράλληλα με την κατάταξη των δεδομένων σε τύπου cross-section ή time-series, θα μπορούσαμε επίσης να τα ξεχωρίσουμε σε continuous ή discrete, με την έννοια ακριβώς που υποδηλώνουν οι ονομασίες τους. Τα Continuous δεδομένα μπορούν να πάρουν οποιαδήποτε αξία και δεν περιορίζονται σε συγκεκριμένα νούμερα. Οι αξίες τους περιορίζονται μόνο από την ακρίβεια. Για παράδειγμα, η απόδοση από την ενοικίαση ενός ακινήτου μπορεί να είναι 6,2%, 6,24%, ή 6,238% κλπ. Αντίθετα, τα discrete δεδομένα μπορούν να πάρουν μόνο συγκεκριμένες αξίες που τις περισσότερες φορές είναι ακέραιοι αριθμοί που συνήθως ορίζονται σαν count numbers. Για παράδειγμα, ο αριθμός των ατόμων σε έναν συγκεκριμένο συρμό Μετρό ή ο αριθμός των μετοχών που διακινήθηκαν σε μια μέρα. Στην περίπτωση αυτή, το να έχεις 86,3 επιβάτες στο Μετρό ή 5857,5 μετοχές δεν έχει νόημα. Το πιο απλό παράδειγμα discrete μεταβλητής είναι μια τυχαία μεταβλητή Bernoulli ή binary random που μπορεί να πάρει μόνο τις αξίες 0 ή 1- για παράδειγμα, αν επανειλημμένα ρίξουμε ένα νόμισμα, μπορούμε να ορίσουμε την κορώνα σαν 1 και τα γράμματα σαν 0.

3.3.5 Cardinal, Ordinal και Nominal αριθμοί.

Ενας άλλος τρόπος που μπορούμε να ταξινομήσουμε αριθμούς είναι με βάση το αν είναι Cardinal, Ordinal και Nominal αριθμοί. Cardinal είναι οι αριθμοί όπου οι πραγματικές αριθμητικές τιμές που μια συγκεκριμένη μεταβλητή παίρνει έχουν νόημα, και όπου υπάρχει μια ίση απόσταση ανάμεσα στις αριθμητικές τιμές. Αντίθετα, οι Ordinal αριθμοί μπορούν να ερμηνευθούν μόνο σαν παρέχοντες μια θέση ή μια διάταξη. Έτσι, για τους Cardinal αριθμούς η αξία 12 σημαίνει ένα μέγεθος που είναι διπλάσιο της αξίας 6. Παραδείγματα Cardinal αριθμών θα ήταν η τιμή μιας μετοχής ή ενός ακινήτου, ή ο αριθμός των ακινήτων σε ένα δρόμο. Αντίθετα, στην κλίμακα των Ordinal το νούμερο 12 μπορεί να θεωρείται «καλύτερο» από το νούμερο 6, αλλά όχι υποχρεωτικά διπλάσια καλύτερο. Σχετικό παράδειγμα θα ήταν η κατάταξη ενός αθλητή σε έναν αγώνα, (η 2^η θέση είναι καλύτερη από την 4^η, αλλά δε θα είχε νόημα να πούμε ότι είναι δύο φορές καλύτερη), ή το επίπεδο που φτάνουμε σε ένα ηλεκτρονικό παιχνίδι.

Ο τελευταίος τύπος δεδομένων που μπορεί να συναντήσουμε είναι η περίπτωση όπου δεν υπάρχει καμιά απολύτως φυσική διάταξη των αξιών, δηλαδή το νούμερο 12 είναι απλά διαφορετικό από το νούμερο 6 αλλά δεν μπορεί με καμιά έννοια να θεωρηθεί καλύτερο ή χειρότερο. Τέτοια δεδομένα εμφανίζονται συχνά όταν κάποιες αριθμητικές αξίες ορίζονται αυθαίρετα, όπως οι τηλεφωνικοί αριθμοί, ή όπως όταν ορίζονται κάποιοι κωδικοί σε ποιοτικά στοιχεία (π.χ, όταν περιγράφουμε την τιμή διαπραγμάτευσης μιας μετοχής στις ΗΠΑ το «1» μπορεί να χρησιμοποιείται για να σημαίνει το NYSE, το «2» το NASDAQ και το «3» το AMEX). Μερικές φορές, τέτοιες μεταβλητές ονομάζονται Nominal. Οι μεταβλητές Cardinal, Ordinal και Nominal πιθανόν να απαιτούν διαφορετικές προσεγγίσεις όσον αφορά τη μοντελοποίηση ή τουλάχιστον διαφορετικό χειρισμό.

3.4 Θεωρία χρονοσειράς

3.4.1 Ορισμός χρονοσειράς

Χρονοσειρά είναι μια σειρά από παρατηρήσεις - observations που λαμβάνονται σε συγκεκριμένες χρονικές στιγμές ή συγκεκριμένες περιόδους και έχουν μεταξύ τους

ιση απόσταση. Πιο συγκεκριμένα, στις χρονοσειρές μελετάται η δυναμική εξέλιξη που προκύπτει από την μεταβλητή πάνω στην οποία γίνεται η ανάλυση. Στις χρονοσειρές η ανάλυση του κάθε χαρακτηριστικού ή της κάθε μεταβλητής γίνεται με σταθερό χρονικό εύρος (sampling time). Φυσικά, υπάρχουν και εξαιρέσεις όπου κάποια προβλήματα δεν έχουν σταθερό sampling time και χρειάζονται διαφορετική προσέγγιση στην ανάλυση της χρονοσειράς. Ένα τέτοιο παράδειγμα, είναι η ανάλυση χρηματιστηριακών τιμών, όπου προκύπτει μεταβλητό sampling time, αφού μεσολαβούν Σαββατοκύριακα ή ακόμα και αργίες.

Ένα βασικό στάδιο κατά τη διάρκεια της ανάλυσης είναι να εκτιμηθεί το εκάστοτε σύστημα που παράγεται από τη χρονοσειρά και φυσικά η δημιουργία προβλέψεων για μελλοντικές τιμές πάνω σε αυτό το σύστημα που δημιουργήθηκε με βάση το χαρακτηριστικό που εξετάζεται. Πριν ξεκινήσει η ανάλυση χρονοσειράς υπάρχουν ορισμένες υποθέσεις που πρέπει να ελεγχθούν, έτσι ώστε τα συμπεράσματα από αυτή να είναι πράγματι ουσιαστικά. Η πρώτη υπόθεση που πρέπει να απορριφθεί είναι ότι η μεταβολή στις τιμές του χαρακτηριστικού που αναλύεται είναι τυχαίες, αυτή η περίπτωση ονομάζεται λευκός θόρυβος (white noise). Σε περίπτωση που οι παρατηρήσεις μας δεν είναι ανεξάρτητες, οι πληροφορίες που προσφέρει η χρονοσειρά μπορεί να δίνονται με διαφορετικούς τρόπους και ως εκ τούτου υπάρχουν ορισμένα χαρακτηριστικά τα οποία θα πρέπει να αναλυθούν πριν την προσαρμογή του μοντέλου στη χρονοσειρά.

Το πρώτο χαρακτηριστικό ονομάζεται στασιμότητα (stationarity). Η στασιμότητα με απλά λόγια σημαίνει ότι οι διακυμάνσεις των τιμών που δημιουργούν την χρονοσειρά δεν διαφέρουν με τη πάροδο του χρόνου. Αντίθετα, μία μη στάσιμη χρονοσειρά περιέχει πολλές φορές διάφορες τάσεις, αλλαγές δηλαδή στις μέσες τιμές με τη πάροδο του χρόνου. Ένα παράδειγμα θα ήταν φυσικά οι αλλαγές στις μέσες τιμές των τιμών ενέργειας παγκοσμίως με την πάροδο του χρόνου. Κάτι άλλο που μπορεί να εμφανίσει μία μη στάσιμη χρονοσειρά είναι στοιχεία περιοδικότητας. Η τελευταία σχετίζεται με τιμές που μπορεί να παίρνει η μεταβλητή που εξετάζεται ανάλογα με συγκεκριμένους μήνες, τρίμηνα ή εποχές γενικότερα (στη τελευταία περίπτωση αναφέρεται και ως εποχικότητα). Ένα παράδειγμα μπορεί να είναι το ποσοστό αποχής από τις εκλογές, όταν εκείνες γίνονται στους καλοκαιρινούς μήνες.

Το δεύτερο χαρακτηριστικό που πρέπει να αναλυθεί ονομάζεται αιτιοκρατία ή στοχαστικότητα. Η συντριπτική πλειοψηφία των χρονοσειρών που προκύπτουν από τον πραγματικό κόσμο, δηλαδή πραγματικές μεταβλητές και δεδομένα περιέχουν ένα μέρος που αντιστοιχεί σε θόρυβο. Έτσι, δε θα ήταν παράλογο να πούμε πως σχεδόν όλες οι χρονοσειρές που βασίζονται σε πραγματικά δεδομένα είναι κατά κάποιο τρόπο στοχαστικές. Ένα από τα μεγαλύτερα προβλήματα που μπορεί να προκύψει κατά τη διάρκεια της ανάλυσης χρονοσειρών είναι να εντοπιστεί το αιτιοκρατικό μέρος που δημιουργεί η χρονοσειρά. Σε περίπτωση που αυτό καταλαμβάνει ένα μικρό μέρος στην δυναμική της χρονοσειράς λένε συνήθως ότι το σύστημα είναι στοχαστικό και τελικά η προσέγγιση βασίζεται στη στατιστική περιγραφή του εκάστοτε συστήματος. Στην αντίθετη περίπτωση, όπου στο σύστημα που δημιουργείται το αιτιοκρατικό στοιχείο είναι κυρίαρχο, τότε πρέπει να αξιοποιηθούν εργαλεία τα οποία είναι κατάλληλα για τέτοιες συνθήκες, όπως είναι η ανίχνευση κυρίων περιόδων.

Ένα ακόμη χαρακτηριστικό που χρίζει ανάλυσης είναι η γραμμικότητα ή η μη γραμμικότητα. Όπως έγινε κατανοητό από τα προηγούμενα, οι δύο αυτές περιπτώσεις σχετίζονται ισχυρά με την αιτιοκρατία. Όμως σαφώς είναι δύο έννοιες που πρέπει να αναλυθούν ξεχωριστά λόγω της βαρύτητας που έχουν. Για ένα σύστημα, γραμμικότητα σημαίνει ότι οι μεταβλητές του έχουν γραμμική αλληλεπίδραση μεταξύ τους. Αυτό σημαίνει πως αν το σύστημα αναλυόταν κάθε όρος του θα ήταν γραμμικός σε σχέση με τις μεταβλητές του. Σε αντίθετη περίπτωση θα λέγαμε πως αυτό το σύστημα είναι μη-γραμμικό. Όσον αφορά τη χρονοσειρά αυτός ο διαχωρισμός έχει σημασία, καθώς στη πρώτη περίπτωση της γραμμικότητας θα πρέπει να οριστεί η δυναμική της χρονοσειράς ως γραμμικός συνδυασμός των παρατηρήσεων που προηγήθηκαν. Αντιθέτως, στην περίπτωση της μη γραμμικότητας η δυναμική ορίζεται ακριβέστερα αν λάβουμε υπόψιν και τη συνδυασμένη επίδραση των παρατηρήσεων που προηγήθηκαν σε διαφορετικά χρονικά σημεία ή στα ίδια.

3.4.2 Βασικά Χαρακτηριστικά Χρονοσειράς

Στη παρούσα υποενότητα μελετώνται κάποια από τα βασικά χαρακτηριστικά των χρονοσειρών.

Στασιμότητα, τάση και περιοδικότητα

Σε περίπτωση όπου η χρονοσειρά εμφανίσει τάση ή περιοδικότητα αυτό σημαίνει πως τα στατιστικά χαρακτηριστικά που προκύπτουν από το σύστημα της εκάστοτε χρονοσειράς μεταβάλλονται με τη πάροδο του χρόνου και άρα η χρονοσειρά δεν είναι στάσιμη. Η αυστηρή στασιμότητα μαθηματικά εκφράζεται ως η διατήρηση στο χρόνο t της ίδιας κατανομής των $\{x_t, x_{t+1}, \dots, x_{t+\tau}\}$ για κάποιο αυθαίρετο παράθυρο υστερήσεων τ . Η συνθήκη στασιμότητας περιορίζεται στο ότι διατηρούνται η μέση τιμή και η αυτοδιασπορά και ονομάζεται συνήθως ασθενής στασιμότητα. Στην ουσία αντί της αυτοδιασποράς εξετάζεται η σταθερότητα της διασποράς με τη πάροδο του χρόνου. Σε περίπτωση μη-στασιμότητας δημιουργείται ένα αρκετά σοβαρό ζήτημα κατά τη διάρκεια της ανάλυσης και αυτό εντείνεται όταν ο στόχος είναι η δημιουργία προβλέψεων. Στις χρονοσειρές εκείνες οι οποίες υποφέρουν από έντονη περιοδικότητα ή εποχικότητα θα ήταν θεμιτό πρώτα οι επιδράσεις αυτές να γίνουν ουδέτερες και μετέπειτα να ξεκινήσει η ανάλυση της. Φυσικά υπάρχουν συγκεκριμένοι στατιστικοί έλεγχοι ώστε να ερευνηθεί η περίπτωση της στασιμότητας σε μία χρονοσειρά.

Απαλοιφή τάσης και περιοδικότητας

Μια χρονοσειρά μπορεί να σπάσει σε τρία συστατικά, με βάση την παρακάτω εξίσωση

$$x_t = \mu_t + s_t + y_t$$

όπου μ_t είναι το συστατικό που δηλώνει την τάση, s_t δηλώνει την περιοδικότητας κάποιας συγκεκριμένης περιόδου και τέλος y_t είναι η χρονοσειρά που αφορά τα υπόλοιπα-residuals, δηλαδή αυτό που μένει αν αφαιρεθεί από την παρατηρούμενη χρονοσειρά η τάση και η περιοδικότητα. Τα δύο τελευταία είναι τα δύο συστατικά τα οποία δεν προσφέρουν κάποια πληροφορία ως προς τη δυναμική του συστήματος,

δηλαδή το πόσο εξαρτάται η παρατήρηση x_t από τις παρατηρήσεις που προηγήθηκαν.

Το τι χρειάζεται να γίνει εξαρτάται κάθε φορά από το πρόβλημα. Πιο συγκεκριμένα, υπάρχουν προβλήματα στα οποία το σύνολο της πληροφορίας της χρονοσειράς περιέχεται στον παράγοντα της τάσης και της περιοδικότητας. Σε αυτές τις περιπτώσεις όλη η προσοχή επικεντρώνεται στον υπολογισμό των μ_t και s_t . Σε άλλα προβλήματα στόχος είναι να αφαιρεθεί η τάση και η περιοδικότητα. Έτσι θα πρέπει αφού υπολογιστούν οι παράγοντες για τα δύο συστατικά, να αφαιρεθούν στην συνέχεια από το x_t , ώστε να δημιουργηθεί η χρονοσειρά των y_t που αντιστοιχούν στα υπόλοιπα όπως αναφέρθηκε ήδη.

Η χρονοσειρά των υπολοίπων δίνεται από την παρακάτω εξίσωση:

$$y_t = x'_t - s_t = x_t - \mu_t - s_t$$

Στην παρακάτω εξίσωση έχει αφαιρεθεί από την χρονοσειρά η τάση και περιοδικότητα.

Υπάρχουν περιπτώσεις που το συστατικό των τάσεων δεν εκφράζεται καταλλήλως ως γνωστή συνάρτηση του χρόνου, αλλά πολλές φορές έχει στοιχεία τυχαιότητας και ως εκ τούτου αναφέρεται ως στοχαστική τάση. Σε περίπτωση που συμβαίνει αυτό, για την απαλοιφή της τάσης προτιμάται ένας άλλος τρόπος που βασίζεται στις πρώτες διαφορές και συγκεκριμένα στην παρακάτω εξίσωση:

$$y_t = \nabla x_t = x_t - x_{t-1}$$

Για την εξάλειψη της περιοδικότητας χρησιμοποιούνται παρόμοιοι τρόποι με αυτούς την εξάλειψης της τάσης, ενώ υπάρχουν και τρόποι που βασίζονται σε άλλους παράγοντες, όπως αυτός που σχετίζεται με τον κινούμενο μέσο και ονομάζεται φίλτρο κινούμενου μέσου. Επειδή αυτές οι τεχνικές λόγω της φύσης της χρονοσειράς που προέκυψε σε αυτήν την διπλωματική εργασία δεν χρησιμοποιήθηκαν, δεν αναλύονται περαιτέρω.

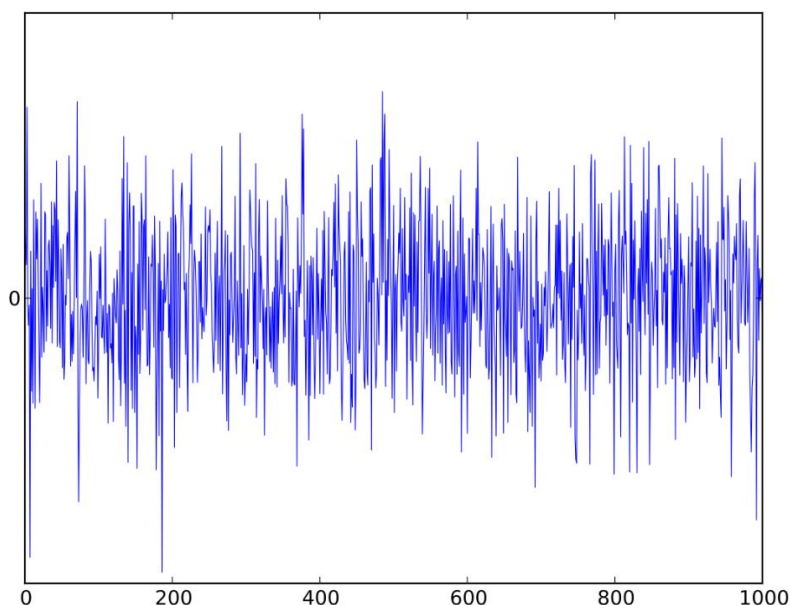
3.4.3 Συσχέτιση σε χρονοσειρά

Στην επόμενη υποενότητα θα αναλυθεί η παρουσία ή μη συσχετίσεων μέσα σε μία χρονοσειρά. Αρχικά, θα αναλυθούν δύο περιπτώσεις όπου η παρουσία συσχέτισης είναι μηδενική. Η μία περίπτωση αντιστοιχεί σε στάσιμη χρονοσειρά και η άλλη σε μη στάσιμη.

3.4.3.1 Συσχέτιση σε χρονοσειρά - Λευκός θόρυβος – White noise

Όσον αφορά τη περίπτωση του λευκού θορύβου, αν θεωρήσουμε τα διαδοχικά στοιχεία της εκάστοτε χρονοσειράς ως τυχαίες μεταβλητές, τότε αυτή λέγεται ότι αποτελείται από ανεξάρτητες τυχαίες μεταβλητές με ίδια κατανομή όταν οι x_t , x_{t+1} , \dots , $x_{t+\tau}$ τυχαίες μεταβλητές για $\tau > 1$ έχουν την ίδια κατανομή και είναι ανεξάρτητες μεταξύ τους. Μια τέτοια χρονοσειρά είναι εντελώς τυχαία και δεν περιέχει καθόλου αυτοσυσχετίσεις, είτε πρόκειται για γραμμικές είτε για μη γραμμικές. Τέτοιου είδους χρονοσειρές λέγονται και λευκού θορύβου. Τέλος, αν τα στοιχεία της χρονοσειράς ακολουθούν κανονική (Gauss) κατανομή, τότε η χρονοσειρά λέγεται Γκαουσιανός λευκός θόρυβος.

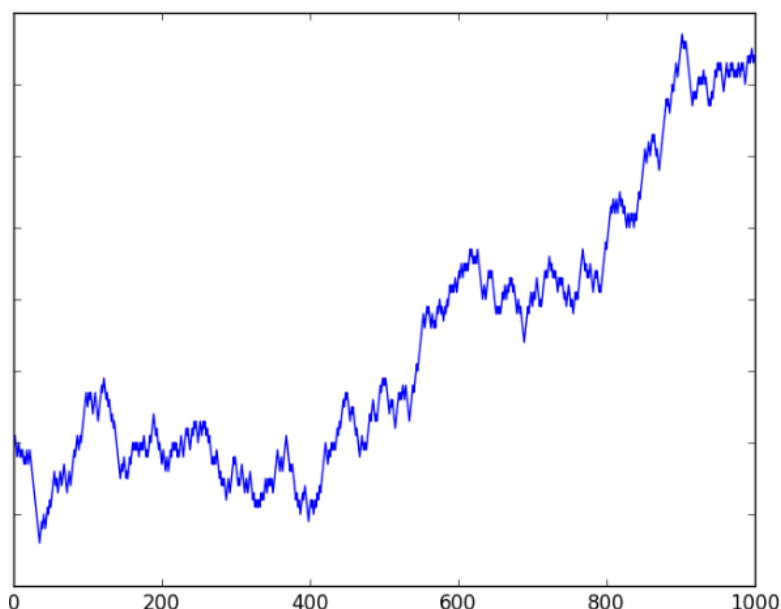
Παράδειγμα γραφήματος Γκαουσιανού λευκού θορύβου



3.4.3.2 Συσχέτιση σε χρονοσειρά - Τυχαίος περίπατος – Random walk

Η περίπτωση του τυχαίου περιπάτου εκφράζει μία μη στάσιμη χρονοσειρά που σημαίνει πως κάθε στοιχείο της χρονοσειράς προκύπτει από την προηγούμενη παρατήρηση και την πρόσθεση σε αυτήν μίας τυχαίας τιμής. Η ονομασία αυτού προκύπτει από το ότι η χρονοσειρά δημιουργείται από την κίνηση κάποιου πάνω σε μια ευθεία γραμμή, που σε κάθε χρονική στιγμή κάνει μία τυχαία κίνηση είτε αυτή είναι προς τα εμπρός είτε προς τα πίσω σε σχέση με το σημείο όπου εκείνος βρίσκεται. Ο τυχαίος περίπατος θα μπορούσε να οριστεί και ως το άθροισμα όλων των τυχαίων αυτών ως προς μία αρχική χρονική στιγμή t . Οι συσχετίσεις που φαίνεται να υπάρχουν σε ένα διάγραμμα ιστορίας τυχαίου περιπάτου προκύπτουν από το ότι η τυχαία κίνηση-βήμα σε κάθε χρονική στιγμή t , έχει γνωστό σημείο εκκίνησης. Αυτό οδηγεί σε στοχαστικές τάσεις και για αυτόν τον λόγο οι χρονοσειρές τυχαίου περιπάτου είναι μη στάσιμες.

Παράδειγμα γραφήματος τυχαίου περιπάτου



3.4.4 Γραμμική συσχέτιση

Συνήθως με τον όρο αυτοσυσχέτιση εννοούμε την γραμμική αυτοσυσχέτιση. Η αυτοσυσχέτιση ρ_τ για κάποια υστέρηση τ είναι ο συντελεστής συσχέτισης δύο σημείων της χρονοσειράς που απέχουν χρονικά τ κινήσεις – βήματα. Οι ορισμοί αυτοσυσχέτισης μπορούν να έχουν νόημα μόνο όταν η χρονοσειρά που δημιουργείται είναι στάσιμη. Όταν αυτή δεν ανήκει στην κατηγορία των στάσιμων χρονοσειρών δε γίνεται η αυτοσυσχέτιση να οριστεί ως συνάρτηση της υστέρησης αλλά ορίζεται για κάθε χρονική στιγμή t . Συνήθως όταν υπολογίζεται η συνάρτηση αυτοσυσχέτισης ως προς την υστέρηση για μια μη στάσιμη χρονοσειρά με τάσεις, έχει υψηλές τιμές και φθίνει πολύ αργά με την υστέρηση. Αυτό δείχνει ότι ισχυρές συσχετίσεις μεταξύ σημείων που βρίσκονται χρονικά κοντά είναι παρούσες. Αντίστοιχα όταν υπολογίζεται η αυτοσυσχέτιση μη στάσιμης χρονοσειράς όπου υπάρχει έντονη η παρουσία της περιοδικότητας ή εποχικότητας τότε συνήθως εμφανίζονται ταλαντώσεις με κορυφές σε υστερήσεις, οι οποίες είναι όμως πολλαπλάσια της περιοδικότητας. Στην ειδική περίπτωση όπου έχουμε χρονοσειρά με τάση ή εποχικότητα, η μορφή της αυτοσυσχέτισης δεν προσφέρει πληροφορίες για τις συσχετίσεις σε κάποιες από τις υστερήσεις. Αυτές τις ελλείπουσες πληροφορίες μπορούμε να τις πάρουμε όμως από την αυτοσυσχέτιση της χρονοσειράς των υπολοίπων. Συγκεκριμένα, αυτό που μας ενδιαφέρει είναι να δούμε αν η χρονοσειρά των υπολοίπων έχει μηδενικές αυτοσυσχετίσεις.

Καταλήγουμε, πως το ερώτημα που προκύπτει, δηλαδή το αν μια στάσιμη χρονοσειρά, που έχουν αφαιρεθεί τα συστατικά της τάσης και της περιοδικότητας είναι ανεξάρτητη, είναι εξαιρετικά σημαντικό κατά τη διάρκεια της ανάλυσης χρονοσειρών και αποτελεί το πρωταρχικό στάδιο ανάλυσης. Πιο συγκεκριμένα, αποτελεί την πρωταρχική μηδενική υπόθεση που πρέπει να απορριφθεί από τον ερευνητή έτσι ώστε να προχωρήσει στην ανάλυση της στάσιμης χρονοσειράς και να καταφέρει μέσω στοχαστικών ή αιτιοκρατικών μοντέλων να εξαγάγει συμπεράσματα από την χρονοσειρά. Για τη μηδενική υπόθεση της ανεξαρτησίας έχουν αναπτυχθεί διάφοροι έλεγχοι, κάποιοι από τους οποίους μέσω της γλώσσας R, εφαρμόστηκαν και στην παρούσα διπλωματική εργασία. Περισσότερα πάνω σε αυτό το θέμα δίνονται στην ενότητα των αποτελεσμάτων.

3.4.5 Θεωρία Μοντέλων ARIMA

Τα στοχαστικά μαθηματικά μοντέλα όπου στοχεύουν στην ανάλυση και την πρόβλεψη της εξέλιξης συγκεκριμένων μεγεθών ονομάζονται ολοκληρωμένα αυτοπαλινδρομικά μοντέλα κινητών μέσων όρων – ARIMA. Σε σχέση με μοντέλα όπως η παλινδρόμηση και συγκεκριμένα η γραμμική παλινδρόμηση, για την οποία πρέπει να υπάρχει η γνώση για τους παράγοντες από τους οποίους εξαρτάται το κάθε μέγεθος, και που η μελλοντική πρόβλεψη τους είναι σχεδόν αδύνατη, τα μοντέλα ARIMA στηρίζονται στην πιθανότητα όπου μια τιμή ενός συγκεκριμένου μεγέθους θα είναι εντός του διαστήματος.

Το μεγαλύτερο μέρος της μελέτης πάνω σε αυτά τα μοντέλα εκπονήθηκε από τους Box και Jenkins και για αυτό πολλές φορές τα μοντέλα αυτά παίρνουν το όνομα τους.

Η πιο γενική δομή ενός μοντέλου ARIMA περιέχει έναν τυχαίο παράγοντα, τις οποίες τιμές για το μέγεθος με βάση τις περιόδους που προηγήθηκαν και τέλος ορισμένους στοχαστικούς παράγοντες και φυσικά το μοντέλο εκφράζεται και ως γραμμικός συνδυασμός των συστατικών του. Στόχος του αναλυτή είναι να βρει ποιος από τους συνδυασμούς είναι ο καταλληλότερος, κάτι που σε αυτή τη περίπτωση έχει να κάνει με το ποιος παρέχει την καλύτερη ποιότητα πρόβλεψης. Ο αναλυτής δεν θα είναι πάντα σίγουρος ότι έχει βρεθεί η βέλτιστη επιλογή, αλλά μπορεί να την προσεγγίζει στην εκάστοτε περίπτωση.

Για να εφαρμόσει κανείς μοντέλα ARIMA, πρέπει να τηρούνται κάποιες προϋποθέσεις. Η πρώτη έχει να κάνει με το ότι η χρονοσειρά θα πρέπει να ανήκει στην κατηγορία των διακριτών, δηλαδή οι παρατηρήσεις της να αφορούν ίσες αποστάσεις μεταξύ των χρονικών σημείων. Η δεύτερη προϋπόθεση έχει να κάνει με το ότι η χρονοσειρά θα πρέπει να είναι στάσιμη και όχι μη στάσιμη, δηλαδή η μέση τιμή και η διακύμανση θα πρέπει να είναι σταθερές σε όλες τις χρονικές στιγμές. Λόγω της στασιμότητας, η χρονοσειρά δεν θα έχει πλέον εξάρτηση στο χρόνο και θα μπορεί να γίνει η στοχαστική ανάλυση της.

Η τελευταία προϋπόθεση σχετίζεται με την δυνατότητα προβλέψεων με μεγάλη ακρίβεια μόνο βραχυπρόθεσμα, καθώς οι προβλέψεις στηρίζονται στις προηγούμενες τιμές. Έτσι γίνεται κατανοητό πως όσο πιο μακροπρόθεσμη γίνεται η

απόπειρα πρόβλεψης, τόσο μειώνονται τα διαθέσιμα δεδομένα και άρα η ακρίβεια και η αξιοπιστία.

Η διαδικασία της επιλογής του καταλληλότερου μοντέλου ARIMA για κάθε χρονοσειρά απαιτεί μελέτη και σίγουρα δεν είναι κάτι προφανές τις περισσότερες φορές. Πολλές φορές κάποια από τα μοντέλα υπάρχει περίπτωση να ταυτιστούν. Αυτό σημαίνει ότι η τελική επιλογή αφήνεται στην κρίση του αναλυτή. Επίσης, ένας παράγοντας που πάντα πρέπει να λαμβάνεται υπόψιν είναι η πολυπλοκότητα του μοντέλου, αφού εάν αυτή είναι μεγάλη τότε το μοντέλο μπορεί να απορριφθεί παρά την καλή προσαρμογή του. Δεν πρέπει να ξεχνάμε ότι η καλή προσαρμογή δεν επιφέρει απαραίτητα και μικρότερα σφάλματα στην πρόβλεψη.

Η διαδικασία με την οποία βρίσκουμε το βέλτιστο μοντέλο ARIMA συνήθως περιλαμβάνει τρία βήματα. Το πρώτο αφορά την αναγνώριση, κατά το στάδιο της οποίας επιλέγεται το μοντέλο ή το σύνολο των μοντέλων ARIMA, τα οποία θεωρούνται από τον αναλυτή κατάλληλα ώστε να μπορούν να περιγράψουν την χρονοσειρά με σωστό τρόπο. Το δεύτερο στάδιο είναι αυτό της εκτίμησης. Στο συγκεκριμένο στάδιο υπολογίζεται για το σύνολο των μοντέλων οι παράμετροι p , d , q και στην συνέχεια υλοποιούνται. Ο πιο συνηθισμένος τρόπος είναι αυτός μέσω της προσδοκώμενης πιθανοφάνειας. Η τελευταία απεικονίζει με βάση τις παραμέτρους το πόσο πιθανό είναι να προσεγγιστούν οι πραγματικές τιμές της κάθε χρονοσειράς. Στόχος δηλαδή είναι να επιλεγούν και να υπολογίζονται τα p, d, q με τέτοιο τρόπο ώστε να μεγιστοποιείται η πιθανοφάνεια. Η επιλογή αυτών των παραμέτρων θυμίζει αρκετά την διαδικασία της γραμμικής παλινδρόμησης, όπου με τις αντίστοιχες παραμέτρους γίνεται προσπάθεια για να ελαχιστοποιηθεί το άθροισμα των τετραγώνων των υπολειπομένων σφαλμάτων. Το τρίτο στάδιο είναι αυτό του διαγνωστικού ελέγχου. Σε αυτό υλοποιούνται στατιστικά τεστ έτσι ώστε να ελεγχθεί το αν και ποια από τα μοντέλα που επελέγησαν είναι στατιστικά σημαντικά. Οι έλεγχοι που επικεντρώνονται στην προσαρμογή των μοντέλων γίνονται πάντα στα αρχικά δεδομένα μελετώντας συγκεκριμένα την κατανομή των σφαλμάτων πρόβλεψης. Αν το μοντέλο είναι “ιδανικό”, τα σφάλματα που θα προκύπτουν θα οφείλονται μόνο σε τυχαία στοιχεία και άρα δεν θα υπάρχουν συσχετίσεις μεταξύ τους. Βέβαια τα σφάλματα ποτέ δεν είναι εντελώς ασυσχέτιστα μεταξύ τους.

Άλλα κριτήρια που βοηθούν στο να επιλέξουμε το καταλληλότερο μοντέλο είναι το Akaike's Information Criterion, το οποίο επελέγη και στην παρούσα διπλωματική εργασία, ενώ υπάρχει και το Bayesian Information Criterion. Εναλλακτικά τα δύο κριτήρια αυτά ονομάζονται AIC και BIC (Brooks, 2018). Τα δύο αυτά κριτήρια λαμβάνουν υπόψιν κυρίως την πολυπλοκότητα του μοντέλου. Πιο συγκεκριμένα, δείχνουν αν η αύξηση της πολυπλοκότητας ενός μοντέλου αυξάνει την πιθανότητα προσέγγισης των πραγματικών τιμών και το κατά πόσο αξίζει αυτό να χρησιμοποιηθεί, έχοντας ως στόχο πάντα την ιδανική περιγραφή της χρονοσειράς από το μοντέλο.

Όσον αφορά τα μοντέλα πρόβλεψης η αύξηση της πολυπλοκότητας μειώνει την προκατάληψη του. Η αύξηση όμως των παραγόντων του εκάστοτε μοντέλου φέρνει αντίστοιχη αύξηση και στην διακύμανση των σφαλμάτων και άρα σε έλλειψη ακρίβειας στις προβλέψεις, ένα φαινόμενο που ονομάζεται και *over-fitting*. Ο κύριος στόχος κατά την επιλογή των παραγόντων είναι λοιπόν να επιτευχθεί μεγάλη ακρίβεια με όσο μικρότερη προκατάληψη γίνεται.

Τα κριτήρια AIC και BIC μας πληροφορούν μόνο για το ποιο από τα μοντέλα που εξετάζονται είναι το καλύτερο. Λόγω έλλειψης υπόθεσης ως βάση τους τα δύο κριτήρια δεν είναι σε θέση να μας πληροφορήσουν για το εάν αυτό το μοντέλο που επελέγη είναι σε θέση να περιγράψει την χρονοσειρά επαρκώς και ικανοποιητικά. Για τον παραπάνω λόγο η τελική επιλογή μοντέλου θα γίνεται με βάση την σύγκριση των διάφορων τιμών από όλα τα υπό εξέταση μοντέλα. Αυτό που ψάχνουμε δηλαδή, είναι το μοντέλο που ελαχιστοποιεί τα κριτήρια με βάση την μέγιστη πιθανοφάνεια.

Ο υπολογισμός του AIC γίνεται με βάση την παρακάτω σχέση:

$$AIC = -2\log L + 2(p + q + k + 1)$$

Με k να είναι ίσο με το μηδέν όταν και η σταθερά του μοντέλου αυτού c είναι ίση με το μηδέν και k ίσο με ένα όταν c διάφορη του μηδενός.

Το κριτήριο BIC είναι αρκετά παρόμοιο με το AIC, όμως δίνει μεγαλύτερο βάρος στο επίπεδο πολυπλοκότητας του εκάστοτε μοντέλου έχοντας ως στόχο να αποφευχθεί

το φαινόμενο over-fitting που αναφέραμε προηγουμένως. Το κριτήριο BIC βασίζεται στην παρακάτω σχέση:

$$\text{BIC} = \text{AIC} + \text{Log}(n)(p + q + k + 1)$$

Μετά την ανάλυση των μοντέλων με βάση τα κριτήρια αξιολόγησης που αναφέραμε αλλά και διάφορα κριτήρια βελτιστοποίησης των παραμέτρων των μοντέλων μπορεί κανείς να επιλέξει το καταλληλότερο μοντέλο ARIMA για να περιγράψει την εκάστοτε χρονοσειρά.

Σημαντικό στην επιλογή του κατάλληλου μοντέλου ARIMA και συγκεκριμένα στο στάδιο της αναγνώρισης, δηλαδή το πρώτο, είναι να γνωρίζει ο αναλυτής τους συντελεστές αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Ο συντελεστής αυτοσυσχέτισης (ACF) προκύπτει από το πηλίκο της αυτοδιακύμανσης προς τη διακύμανση της χρονοσειράς. Ο συντελεστής αυτός δηλαδή απεικονίζει το πόσο η τιμή που παίρνει μία χρονοσειρά σε μία περίοδο t εξαρτάται από την τιμή που είχε πάρει η χρονοσειρά k περιόδους πίσω ($t-k$). Οι τιμές που μπορεί να πάρει βρίσκονται στο εύρος -1 με $+1$ και φυσικά δηλώνουν το ύψος της συσχέτισης και το πρόσημο, αν αυτή είναι θετική ή αρνητική. Σε περίπτωση που το αποτέλεσμα του συντελεστή είναι μηδέν τότε δεν υπάρχει καμία συσχέτιση μεταξύ των τιμών των δύο περιόδων.

Ο συντελεστής μερικής αυτοσυσχέτισης με τη σειρά του απεικονίζει και αυτός το πόσο η τιμή της χρονοσειράς για μία περίοδο t επηρεάζεται από την τιμή της χρονοσειράς k περιόδους πίσω χωρίς όμως εδώ να λαμβάνονται υπόψιν οι επίδρασεις που μπορούν να δημιουργούν οι τιμές των παρατηρήσεων μεταξύ τους. Συνήθως ο PACF υπολογίζεται μέσω συντελεστών ελαχίστων τετραγώνων. Σε περίπτωση φυσικά όπου το k είναι ίσο με ένα τότε οι δείκτες των ACF και PACF είναι ταυτόσημοι. Οι συντελεστές των ACF και PACF για τα μοντέλα ARIMA είναι στην ουσία εκτιμητές των πραγματικών συντελεστών της χρονοσειράς, που σημαίνει ότι προσεγγίζουν τις αληθινές τιμές αλλά δεν ταυτίζονται με αυτές.

Όπως έχει ήδη αναφερθεί προϋπόθεση για τα μοντέλα ARIMA είναι η χρονοσειρά να είναι στάσιμη. Για να συμβαίνει αυτό θα πρέπει παράγοντες όπως η μέση τιμή της, η διακύμανση και η συνάρτηση αυτοσυσχέτισης της να παραμένουν σταθερές στο χρόνο. Για να τηρηθεί αυτή η προϋπόθεση πολλές φορές χρειάζονται τροποποιήσεις

στην χρονοσειρά. Μία λύση για την μετατροπή της χρονοσειράς σε στάσιμη είναι να την λογαριθμήσουμε, καθώς μειώνεται η διακύμανση και αυτό είναι κάτι που προσδίδει μεγαλύτερα επίπεδα σταθερότητας. Συγκεκριμένοι κανόνες για το πότε πρέπει να υιοθετείται αυτό το μέτρο δεν υπάρχουν, όμως μεγάλες διακυμάνσεις ή χαμηλά επίπεδα ακρίβειας μπορούν να λειτουργήσουν ως δείκτες για την χρήση του. Άλλες λύσεις είναι σαφώς και οι μετασχηματισμοί Box Cox, έχοντας πάντα ως απαίτηση θετικές τιμές για τις παρατηρήσεις. Μια άλλη λύση είναι η διαφόριση, κυρίως όταν παρατηρείται τάση ή εποχικότητα. Αυτή δημιουργεί μία χρονοσειρά με σταθερό επίπεδο και διακύμανση, αφού στοχεύει στον να μειώνει τις διακυμάνσεις βγάζοντας κομμάτια τάσης ή εποχικότητας από την χρονοσειρά. Είναι σημαντικό να ξεκαθαριστεί πως η χρήση της διαφόρισης δεν πρέπει να είναι υπέρμετρη (υπερδιαφόριση), καθώς αυτό θα μείωνε πολύ τις διαθέσιμες παρατηρήσεις και έτσι η μελέτη εκείνων των χρονοσειρών που δεν μπορούν να χαρακτηριστούν ως μεγάλες θα επηρεάζονταν με πολύ αρνητικό τρόπο. Ένας άλλος παράγοντας που πρέπει να ληφθεί υπόψιν είναι πως χρησιμοποιώντας την διαφόριση μειώνεται η αυτοσυσχέτιση και άρα αυξάνονται οι τυχαίες επιδράσεις. Αυτό ακολούθως μπορεί να δημιουργήσει την απαίτηση για περισσότερο πολύπλοκα μοντέλα, που σαφώς περιπλέκουν την ανάλυση.

3.4.6. Μοντέλα Αυτοπαλινδρόμησης - Autoregressive models- AR(p)

Τα αυτοπαλινδρομικά μοντέλα μελετούν τις γραμμικές σχέσεις που προκύπτουν μεταξύ της τιμής μίας παρατήρησης της χρονοσειράς που ερευνάται και των προηγούμενων τιμών της, σε αντίθεση με τα απλά παλινδρομικά μοντέλα που ερευνούν τις σχέσεις μεταξύ των τιμών της εκάστοτε χρονοσειράς με διάφορους παράγοντες-μεταβλητές που μπορούν να τις επηρεάζουν. Για την περίπτωση της p τάξης το αυτοπαλινδρομικό μοντέλο γράφεται ως εξής:

$$y_t = c + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \dots + \varphi_p y_{t-p} + e_t$$

Στην παραπάνω εξίσωση το φ_i συμβολίζει τους συντελεστές αυτοσυσχέτισης για ένα μοντέλο AR έχοντας ως υστέρηση i και σταθερά c . Η εξίσωση αυτή πρακτικά σημαίνει πως η παρατήρηση y_t επηρεάζεται κατά φ_1 από την ακριβώς προηγούμενη

παρατήρηση, κατά φ_2 από την προ-προηγούμενη και κατά φ_p από την παρατήρηση που προηγήθηκε πριν p περιόδους. Εναλλακτικά για AR(p) και συγκεκριμένα για τη σταθερά, με μ να συμβολίζει τη μέση τιμή, δημιουργείται η παρακάτω εξίσωση:

$$c = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$$

3.4.7. Μοντέλα Κινητού Μέσου Όρου - Moving Average MA(q)

Τα μοντέλα κινητού μέσου όρου απεικονίζουν τυχούσες γραμμικές σχέσεις μεταξύ της παρατήρησης της χρονοσειράς που αναλύεται και των σφαλμάτων που εμφάνισε το εκάστοτε μοντέλο MA σε περιόδους που έχουν προηγηθεί. Ένα μοντέλο MA μπορεί να περιγραφεί αλγεβρικά με τον παρακάτω τρόπο.

$$y_t = c - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \dots - \theta_p e_{t-p} + e_t$$

Στην παραπάνω εξίσωση το θ_i συμβολίζει τους συντελεστές μερικής αυτοσυσχέτισης για ένα μοντέλο MA έχοντας ως υστέρηση i . Η ερμηνεία της εξίσωσης αυτής γίνεται με τον ίδιο τρόπο που έγινε και στα μοντέλα AR(p). Για τα μοντέλα MA τιμή της σταθεράς δίνεται παρακάτω

$$c = \mu$$

Συνδυάζοντας τα μοντέλα AR και MA, παρατηρούμε πως όλα τα μοντέλα AR, μπορούν να γραφτούν ως μοντέλα MA άπειρων όρων.

3.4.8 Μοντέλα ARIMA (p,d,q)

Τα μοντέλα που παρουσιάστηκαν παραπάνω συνδυάζονται και δημιουργούν τα μοντέλα ARIMA που έχουν φυσικά και αυτά ως στόχο την ανάλυση αλλά και την εκπόνηση προβλέψεων πάνω σε στάσιμες χρονοσειρές. Αυτά προκύπτουν από την εισαγωγή στις παραπάνω εξισώσεις της διαφορίσης, η οποία όπως επισημάνθηκε ήδη, στοχεύει στην επίτευξη στάσιμων χρονοσειρών. Πιο συγκεκριμένα τα μοντέλα ARIMA έχουν τάξη p,d,q άρα γράφονται ως ARIMA(p,d,q). Τα μοντέλα ARIMA περιγράφονται αλγεβρικά με τον παρακάτω τρόπο:

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p)(1 - B)^d(1 - B^m)^N y_t = c + (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) e_t$$

Όπως γίνεται κατανοητό το αριστερό μέρος της εξίσωσης αντιστοιχεί στο μοντέλο AR, ο μεσαίος όρος αντιπροσωπεύει τη διαφόριση και τέλος το δεξί μέρος αφορά το μοντέλο MA. Για τη σταθερά και για $n = N = 0$ προκύπτει:

$$c = \mu(1 - \varphi_1 - \varphi_2 - \dots - \varphi_p)$$

Σε αντίθετη περίπτωση η σταθερά ισούται με 0. Τις περισσότερες φορές ο μέσος όρος των διαφορήσιμων-στάσιμων χρονοσειρών είναι ίσος με μηδέν. Στην ουσία όμως οι χρονοσειρές σπάνια είναι απόλυτα στάσιμες κοντά στο 0 και για αυτό το λόγο ακόμα και σε σειρές που έχει ήδη χρησιμοποιηθεί διαφόριση εισάγεται σταθερά c . Στη περίπτωση όπου δεν έχει πραγματοποιηθεί διαφόριση, δηλαδή όταν d είναι ίσο με μηδέν, λέμε ότι αυτό είναι ένδειξη σταθερότητας για την εκάστοτε χρονοσειρά, αλλά αυτή μπορεί να μην είναι αρκετή. Αυτό σημαίνει πως ο εμπλουτισμός του μοντέλου με μία σταθερά είναι αναγκαίος και ίσως να βοηθά στον εντοπισμό του επιπέδου της παραγόμενης χρονοσειράς. Όταν έχει προκύψει διαφόριση πρώτης τάξης, δηλαδή d ίσο με ένα, σημαίνει ότι υπάρχει σταθερή τάση στην εκάστοτε χρονοσειρά, η οποία έχει απαλείφει μετά την πραγματοποίηση της διαφόρισης. Φυσικά πάντα υπάρχει περίπτωση να χρειάζεται να επιβεβαιωθεί το επίπεδο της χρονοσειράς, και εάν συμβαίνει αυτό τότε η εισαγωγή της c είναι απαραίτητη. Αυτή η λύση δεν προτιμάται συνήθως από τους αναλυτές αλλά η επιλογή τις περισσότερες φορές γίνεται μέσω των κριτηρίων AIC και BIC. Τέλος, σε περίπτωση που υπάρχει διαφόριση δεύτερης τάξης, δηλαδή d ίσο με δύο, σημαίνει πως στην χρονοσειρά που ερευνάται εντοπίζεται τάση μέσα στην τάση και ο εμπλουτισμός με σταθερά δεν είναι η κατάλληλη λύση.

Καταλήγοντας, τα μοντέλα ARIMA μπορούν να χρησιμοποιηθούν μόνο σε εκείνες τις χρονοσειρές οι οποίες μπορούν να χαρακτηριστούν ως στάσιμες. Με την στασιμότητα αυτή διασφαλίζονται καλύτεροι και ακριβέστεροι εκτιμητές της μέσης τιμής, της διακύμανσης και φυσικά της συνάρτησης που αφορά την αυτοσυσχέτιση. Το αν η χρονοσειρά είναι στάσιμη μπορεί να ελεγχθεί μέσω ορισμένων περιορισμών-ορίων που επικεντρώνονται στους παράγοντες θ_i και φ_i , οι οποίοι περιεγράφηκαν παραπάνω.

3.4.9 Πρόβλεψη με μοντέλα ARIMA

Όταν ο στόχος είναι ο υπολογισμός μιας τιμής για την χρονοσειρά y για μία χρονική στιγμή-περίοδο t , υπάρχουν ορισμένες πληροφορίες οι οποίες είναι απαραίτητες για την επίτευξη του. Αρχικά απαιτείται η γνώση των τιμών $y_{t-1} \dots y_{t-p}$ και $e_{t-1} \dots e_{t-q}$. Έχοντας ως T σύμβολο για τον ορίζοντα της πρόβλεψης, δημιουργείται ακόμα η απαίτηση για γνώση των $y_{t+T-1} \dots y_{t+T-n}$ και $e_{t+T-1} \dots e_{t+T-n}$. Για παράδειγμα για μία πρόβλεψη τριών περιόδων μπροστά θα πρέπει να υπάρχει γνώση για τρεις περιόδους των y και e ($t, t+1, t+2$). Το πρώτο πρόβλημα που δημιουργείται με αυτές τις απαιτήσεις είναι ότι δεν είναι διαθέσιμες οι τιμές για την δεύτερη πρόβλεψη-περίοδο, δηλαδή οι y_{t+1} και e_{t+1} . Αυτό το πρόβλημα λύνεται εάν επιλεγεί αντί της y_{t+1} η προηγούμενη εκτιμημένη τιμή από την χρονοσειρά, έχοντας σφάλμα ίσο με μηδέν. Αντίστοιχο πρόβλημα και λύση συναντάμε και στην τρίτη περίοδο. Από τα παραπάνω γίνεται κατανοητός και ο λόγος για τον οποίο τα μοντέλα ARIMA υιοθετούνται κυρίως για βραχυχρόνιες προβλέψεις, αφού στην ουσία όσο μεγαλώνει ο χρόνος της πρόβλεψης το ARIMA μετατρέπεται περισσότερο σε απλό μοντέλο AR.

Καταλήγοντας, μέσα από την περιγραφή της θεωρίας των μοντέλων ARIMA έγινε κατανοητό πως αυτά εάν αξιοποιηθούν καταλλήλως, μπορούν να λειτουργήσουν ως ένα πολύ χρήσιμο εργαλείο για την ανάλυση της εξέλιξης των εκάστοτε φυσικών μεγεθών, προσφέροντας δυνατότητες πρόβλεψης. Η προσέγγιση των χρονοσειρών είναι στοχαστική και βασίζονται κυρίως στις παρελθούσες τιμές και πολύ περισσότερο στις πιο κοντινές. Αυτός είναι και ο λόγος που μέσω αυτών οι αναλυτές επικεντρώνονται κυρίως σε βραχυπρόθεσμες προβλέψεις.

Μία από τις βασικές προϋποθέσεις για την σωστή υιοθέτηση του μοντέλου είναι να μην δημιουργεί αχρείασθή πολυπλοκότητα. Αυτό σημαίνει ότι πρέπει να βρεθεί εκείνο το σημείο ισορροπίας, όπου το μοντέλο περιγράφει με τον καλύτερο τρόπο τη χρονοσειρά χωρίς να ξεπερνάει όμως τα όρια της μη θεμιτής πολυπλοκότητας, της εισαγωγής δηλαδή μεγάλου αριθμού συντελεστών. Για να γίνει αυτό αξιοποιούνται τα εργαλεία και τα κριτήρια που αναφέρθηκαν παραπάνω. Εκτός όμως από τα όρια της πολυπλοκότητας πρέπει να τηρείται και η στασιμότητα για την εκάστοτε χρονοσειρά, καθώς μέσω αυτής ο αναλυτής μπορεί να οδηγηθεί σε ακριβείς

εκτιμήσεις των συντελεστών, κάτι που φυσικά είναι προτιμητέο. Συνεπώς ένα μοντέλο ARIMA όταν τηρεί την στασιμότητα και την ισορροπία της πολυπλοκότητας είναι σε θέση να προσφέρει ακριβώς αυτό στο οποίο στοχεύει ο αναλυτής, δηλαδή την ακριβή περιγραφή της χρονοσειράς και προβλέψεις υψηλής ποιότητας για το κοντινό μέλλον.

4.Ανάλυση Μεταβλητών και Μεθοδολογιών

4.1 Εξαρτημένη Μεταβλητή

Όσον αφορά την εξαρτημένη μεταβλητή, αυτή δηλαδή που επικεντρώνεται στα επίπεδα των άμεσων ξένων επενδύσεων, θα χρησιμοποιηθεί μία μέθοδος που υιοθετήθηκε από τους Lensink and Morrissey (2006) αλλά και τον Ahlquist (2006). Η μεταβλητή αυτή αποτελείται από τις καθαρές εισροές άμεσων ξένων επενδύσεων ως ποσοστό του ΑΕΠ κάθε χώρας και αφορά την περίοδο 2003-2020. Τα δεδομένα για αυτή τη μεταβλητή αντλούνται από την βάση δεδομένων της Παγκόσμιας Τράπεζας. Πιο συγκεκριμένα, η βάση παρέχει έναν δείκτη που απεικονίζει τις νέες επενδυτικές εισροές μείον την από-επένδυση για κάθε χώρα από ξένους επενδυτές και αυτό διαιρείται από το ΑΕΠ της. Η μεταβλητή θα ονομαστεί FDI.

4.2 Κύρια Ανεξάρτητη Μεταβλητή

Ως κύρια ανεξάρτητη μεταβλητή θα υιοθετηθεί η Cpi, η οποία θα περιγράφει την διαφθορά. Όσον αφορά την Cpi, θα χρησιμοποιηθεί ένας από τους πιο γνωστούς δείκτες που σχετίζονται με την διαφθορά. Πιο συγκεκριμένα η κύρια ανεξάρτητη μεταβλητή θα βασιστεί στον Corruption Perceptions Index της βάσης δεδομένων του οργανισμού Transparency International. Ο συγκεκριμένος δείκτης βαθμολογεί και κατατάσσει χώρες ή περιοχές ανάλογα με τον επίπεδο διαφθοράς στον δημόσιο τομέα, σύμφωνα με δηλώσεις ειδικών αλλά και ανθρώπων της αγοράς. Το εύρος των τιμών του δείκτη είναι από το 0 για καθόλου διαφθορά έως 100 για απόλυτη διαφθορά στην χώρα.

4.3 Εξαρτημένες Μεταβλητές Ελέγχου

Μεταβλητές που αφορούν την ποιότητα των θεσμών αλλά και την πολιτική σταθερότητα/αστάθεια επελέγησαν για να εισαχθούν στο αρχικό μοντέλο.

Συγκεκριμένα, δημιουργείται η μεταβλητή ROL, η οποία βασίζεται σε δεδομένα από τους δείκτες World Governance της Παγκόσμιας Τράπεζας και για την ακρίβεια στον δείκτη “rule of law”. Ο συγκεκριμένος δείκτης έχει χρησιμοποιηθεί και από τους Alquist et al. (2018). Ο δείκτης αυτός μετράει τα επίπεδα εμπιστοσύνης των μελών της κοινωνίας σε θεσμικά θέματα όπως η εφαρμογή των συμβολαίων, δικαιώματα ιδιοκτησίας, η αστυνομία, τα δικαστήρια αλλά και η πιθανότητα για εγκλήματα και βία. Η αξιολόγηση κάθε χώρας βασίζεται σε μονάδες μίας κανονικής κατανομής με εύρος -2,5 με 2,5 , όπου οι τιμές κοντά στο άνω όριο αποτελούν ένδειξη καλύτερης ποιότητας θεσμών για την συγκεκριμένη χώρα.

Όπως αναλύθηκε και από τους Lensink et al. (2000), το μικρό πολιτικό ρίσκο είναι εξαιρετικά σημαντικός πόλος για την άντληση κεφαλαίων από το εξωτερικό. Δύο δείκτες που σχετίζονται με το πολιτικό ρίσκο, ενσωματώθηκαν στο αρχικό μοντέλο έρευνας της σχέσης άμεσων ξένων επενδύσεων και διαφθοράς ως μεταβλητές ελέγχου.

Η πρώτη από τις δύο μεταβλητές ονομάστηκε POLSTAB και βασίστηκε σε δεδομένα από τους εκτιμητές Political Stability and Absence of Violence/Terrorism από την βάση δεδομένων της Παγκόσμιας Τράπεζας και συγκεκριμένα τους δείκτες Worldwide Governance. Ο δείκτης αυτός μετρά την πιθανότητα πολιτικής αστάθειας, πολιτικά υποκινούμενης βίας ή ακόμα και τρομοκρατικών χτυπημάτων. Οι τιμές που παίρνει αυτή η μεταβλητή-δείκτης είναι -2,5 με 2,5 με τις υψηλότερες τιμές να υποδηλώνουν μεγαλύτερη πολιτική σταθερότητα και χαμηλότερα επίπεδα βίας για την κάθε χώρα που ερευνάται.

Όσον αφορά τον δεύτερο δείκτη που αφορά τον πολιτικό κίνδυνο δημιουργήθηκε η μεταβλητή VOICE. Αυτή βασίστηκε στον εκτιμητή Voice and Accountability του World

Governance Indicators. Ο δείκτης λαμβάνει υπόψιν τις ελευθερίες έκφρασης, είτε πρόκειται για τις προσωπικές είτε για ελευθερία έκφρασης των Μέσων μαζικής ενημέρωσης. Φυσικά βαθμολογείται και η ελευθερία στην επιλογή κυβερνήσεων, δηλαδή κατά τη περίοδο εκλογών. Όπως και στις μεταβλητές ελέγχου που αναφέρθηκαν η βαθμολόγηση κάθε κράτους βασίζεται σε μονάδες κανονικής κατανομής με εύρος -2,5 με 2,5 όπου οι άνω τιμές φανερώνουν ένα κράτος που λειτουργεί πιο δημοκρατικά από ένα κράτος με χαμηλές τιμές του δείκτη.

Για την ανάλυση της επίδρασης του ανθρώπινου κεφαλαίου στις άμεσες ξένες επενδύσεις χρησιμοποιήθηκε η μεταβλητή HUMC. Τα δεδομένα για τον δείκτη αυτόν αντλήθηκαν από το πρόγραμμα ανάπτυξης των Ηνωμένων Εθνών (United Nations Development Program). Τα δεδομένα όπως και για τις άλλες μεταβλητές είναι σε εθνικό επίπεδο και η δημιουργία του δείκτη βασίστηκε πάνω στα επίπεδα εκπαίδευσης, υγείας, αλλά και στο επίπεδο διαβίωσης των πολιτών για κάθε χώρα.

Στο μοντέλο εισάγονται και μεταβλητές σχετικές με τα χρηματοοικονομικά δεδομένα των χωρών και τις επιδόσεις τους.

Αρχικά θα εισαχθεί μία μεταβλητή που θα απεικονίζει το πόσο ανοιχτή είναι μια χώρα στις ξένες επενδύσεις και στο διεθνές εμπόριο, παράγοντες οι οποίοι σύμφωνα με την βιβλιογραφία είναι εξαιρετικά σημαντικοί για την άντληση ξένων κεφαλαίων. Η ονομασία της θα είναι OPTRADE. Η μεταβλητή αυτή θα βασιστεί σε έναν από τους δείκτες που παρέχει η παγκόσμια τράπεζα και συγκεκριμένα στο σύνολο των εισαγωγών και των εξαγωγών υπηρεσιών και προϊόντων ως ποσοστό του ΑΕΠ. Ο δείκτης αυτός χρησιμοποιήθηκε και από τους Lensink και Morrissey (2005). Αφού πρόκειται λοιπόν για ποσοστό, γίνεται κατανοητό πως ο δείκτης αυτός παίρνει μόνο θετικές τιμές. Ο παγκόσμιος μέσος όρος είναι στο 60%

Η δεύτερη μεταβλητή ελέγχου που σχετίζεται με μακροοικονομικά δεδομένα και θα εισαχθεί στα παλινδρομικά μοντέλα θα ονομαστεί GDPPC. Αυτή θα απεικονίζει το κατά κεφαλήν ΑΕΠ ανά χώρα. Η πληροφορία για την μεταβλητή αυτή θα αντληθεί από τις βάσεις δεδομένων της Παγκόσμιας Τράπεζας. Ο δείκτης υπολογίζεται ως το

ΑΕΠ διαιρούμενο από τον πληθυσμό κάθε χώρας στα μέσα κάθε έτους. Όπως αναφέρθηκε και από τον Ahlquist (2006), ο δείκτης αυτός μπορεί υπό προϋποθέσεις να παράξει στατιστικά σημαντικά αποτελέσματα όσον αφορά την κατανομή ξένων κεφαλαίων. Τα δεδομένα δίνονται σε constant 2010 U.S. dollars.

Σύμφωνα με τη βιβλιογραφία σημαντικοί παράγοντες για την κατανομή των πόρων των επενδυτών και συγκεκριμένα των άμεσων ξένων επενδύσεων είναι και το πόσες φορές έχει χρεοκοπήσει μία χώρα αλλά και ο κίνδυνος για μία πιθανή μελλοντική της χρεοκοπία. Για αυτό το λόγο είναι απαραίτητο να ενσωματωθούν στα μοντέλα μεταβλητές που θα αντιπροσωπεύουν δείκτες που σχετίζονται με τα χρέη.

Η πρώτη θα ονομαστεί DEBTGDP και θα απεικονίζει το κρατικό χρέος ως προς το ΑΕΠ της χώρας. Ο παράγοντας αυτός έχει ερευνηθεί και στην εργασία “Capital flight and political risk” από τους Lensink et al. (2000). Η πηγή για τον δείκτη αυτόν θα είναι η βάση δεδομένων του Διεθνούς Νομισματικού Ταμείου.

Table 1: Variable Descriptive Statistics

Variable	Description	Source	Average	Min	Max
FDI	FDI net inflows as a percent of GDP	World Bank	5.4	-41.1	54.6
Cpi	Corruption Perceptions Index	Transparency international	57.0	27.0	80.0
ROL	Inst.Quality/ Rule of Law index	World Bank Governance Ind.	4.9	-41.1	54.6
POLSTAB	Political Stability Indicator	World Bank Governance Ind.	0.3	-2	1,2
VOICE	Voice and Accountability estimate	World Bank Governance Ind.	0.4	-1.8	1.2
OPTRADE	Opemess to Trade	World Bank Development Ind.	113.3	56.2	190.4
GDPPC	GDP divided by midyear population	World Bank Development Ind.	10.1	6.8	27.5
HUMC	Human capital index	UN Development Program	0.8	0.6	0.9

DETBGDP	Governmentnt debt to GDP	IMF Database	40.2	3.8	84.7
IND	Industry Value Added as % of GDP	Transparency international	26.2	14.3	38.8
MAN	Manufacturing Value Added as a % of GDP	Transparency international	15.3	3.7	28.6
INDMM	Industry minus Manufacturing	Transparency international	11.0	5.5	20.7

The data above concern the main variables of the first part of the analysis and data are gathered from various sources. Column 1 presents the name of the variables. Column 2 offers a short description. Column 3 presents the source of the variables. Lastly, column 4 gives the average, column 5 the minimum value and column 6 the maximum value

4.4 Κύριο Συμβατικό Μοντέλο

Στην ενότητα αυτή παρουσιάζεται ο σχηματισμός του μοντέλου. Η βιβλιογραφία δεν παρέχει ένα ενιαίο μοντέλο που να έχει αποδειχθεί ότι περιλαμβάνει όλες τις μεταβλητές που είναι απαραίτητες για να κρατηθούν σταθερές ώστε να απαντήσουν στην υπόθεσή μας. Σε αυτό το μοντέλο, προτείνω τις μεταβλητές ελέγχου που αναφέρθηκαν παραπάνω και συνδέονται με την πολιτική σταθερότητα, τη θεσμική ποιότητα, την ανάπτυξη του ανθρώπινου κεφαλαίου και τα μακροοικονομικά δεδομένα. :

$$FDI_{it} = a + \beta_1 Cpi_{it} + e_{it} \quad (1)$$

Όπου το FDI αντιπροσωπεύει τις άμεσες ξένες επενδύσεις και το Cpi είναι η μεταβλητή που αντιπροσωπεύει τον δείκτη της διαφθοράς.

Μετά τη συμπερίληψη των μεταβλητών ελέγχου, το δεύτερο μοντέλο παλινδρόμησης μπορεί να γραφτεί ως εξής:

$$FDI_{it} = a + \beta_1 Cpi_{it} + \beta_2 M_{it} + \beta_3 Z_{it} + \beta_4 I_{it} + \mu_i + e_{it} \quad (2)$$

όπου μ_i είναι το country fixed effect. Το σύνολο των μεταβλητών M συνδέεται με τον πολιτικό ρίσκο, το κράτος δικαίου και τη θεσμική ποιότητα. Τα διανύσματα Z και I είναι σύνολα μεταβλητών που περιλαμβάνονται στις παλινδρομήσεις, όπως

μεταβλητές που συνδέονται με χρηματοοικονομικές επιδόσεις και μακροοικονομικά δεδομένα για το Z και μεταβλητές που περιγράφουν την εξέλιξη του ανθρώπινου κεφαλαίου για το I. Η δομή του συνόλου δεδομένων είναι unbalanced panel.

Δεδομένου ότι δεν υπάρχει λόγος να υποθέσουμε ότι δεν υπάρχει ετερογένεια, αυτά τα μοντέλα συνήθως χρησιμοποιούν panel fixed effects. Συμπεριλαμβάνοντας fixed effects, ελέγχω τις μέσες διαφορές μεταξύ των χωρών σε τυχόν παρατηρήσιμους ή μη παρατηρήσιμους predictors. Ένα από τα κύρια χαρακτηριστικά της Ανατολικής Ευρώπης είναι οι διαφορές μεταξύ των χωρών.

Πιο συγκεκριμένα, ορισμένες χώρες αποτελούν μέρος της Ευρωπαϊκής Ένωσης, ενώ άλλες όχι. Κάποιοι είναι στην Ευρωζώνη, ενώ άλλοι όχι. Επιπλέον, υπάρχουν σημαντικές διαφορές στα επίπεδα πολιτισμού, ανθρώπινου κεφαλαίου και εξευρωπαϊσμού. Οι συντελεστές σταθερού αποτελέσματος θα απορροφήσουν όλες αυτές τις διαφορές.

Δεδομένου ότι ο αριθμός των οντοτήτων-χωρών δεν είναι πολύ μεγάλος, γεγονός που θα απαιτούσε τη συμπερίληψη μεγάλου αριθμού μεταβλητών dummies, μπορούμε εύκολα να χρησιμοποιήσουμε τόσο το LSDV όσο και τον within estimator για να διαμορφώσουμε τον εκτιμητή. Καταλήγουμε στην πρώτη υπόθεση της εργασίας:

Για H1: $\beta_1 > 0$ σημαίνει πως χαμηλότερη διαφθορά οδηγεί σε μεγαλύτερες εισροές FDI. (Σε περίπτωση στατιστικά σημαντικών αποτελεσμάτων)

5.Εναλλακτική Προσέγγιση: Διεφθαρμένοι Κλάδοι και Επενδύσεις.

5.1 Περιγραφή

Σύμφωνα με τον οργανισμό Transparency International και την διεθνή έρευνα που διεξήγαγε σχετικά με τη διαφθορά και τις δωροδοκίες, η οποία πραγματοποιήθηκε μέσω αναλυτικών συνεντεύξεων με λογιστές, δικηγόρους, τραπεζίτες, κρατικούς λειτουργούς υπουργείων αλλά και στελέχη επιχειρήσεων, συγκεκριμένοι κλάδοι της οικονομίας φαίνεται να είναι πιο επιρρεπής στη διαφθορά.

Μια από τις ερωτήσεις που έγιναν κατά τη διάρκεια των συνεντεύξεων και ίσως η σημαντικότερη ήταν σε ποιους κλάδους συναντούν ή θα περίμεναν να συναντήσουν οι ερωτώμενοι παράνομες πράξεις, όπως οι δωροδοκίες. Τα αποτελέσματα είναι

αρκετά αναμενόμενα και φαίνεται να επιβεβαιώνουν την γενικότερη άποψη που υπάρχει στην κοινωνία, για το ποιοι τομείς της οικονομίας είναι πιο πιθανό να είναι έχουν υψηλότερη συχνότητα τέτοιων συμπεριφορών. Ο τομέας που σύμφωνα με τους ερωτώμενους είναι ο πιο διεφθαρμένος από όλους φαίνεται να είναι ο κατασκευαστικός, η κυριότερη μορφή διαφθοράς σε αυτόν το τομέα όπως έδειξαν οι απαντήσεις είναι η ανάγκη-απαίτηση για δωροδοκίες. Ομοίως, ο τομέας που βρίσκεται στη δεύτερη θέση της λίστας είναι αυτός της άμυνας. Ο χώρος των εξοπλισμών θεωρείται ότι επηρεάζεται κυρίως από τις διμερείς σχέσεις χωρών και την γεωπολιτική γενικότερα, όμως οι απαντήσεις των ερωτώμενων δείχνουν ότι ο τομέας αυτός πάσχει και από δωροδοκίες και διαφθορά, επιβεβαιώνοντας την γενική άποψη ότι η ύπαρξη κρυφής ατζέντας στον τομέα αυτόν είναι κάτι συνηθισμένο. Ο τρίτος πιο διεφθαρμένος τομέας όπως αποκαλύπτει η έρευνα είναι αυτός του πετρελαίου και της ενέργειας γενικότερα. Σε αυτή την περίπτωση, ο ενεργειακός τομέας έχει ήδη χτίσει “κακό όνομα”, κυρίως λόγω των περιβαλλοντικών προβλημάτων που έχει προκαλέσει στο παρελθόν. Όμως από ότι φαίνεται παράνομες συμπεριφορές όπως οι δωροδοκίες βρίσκουν και αυτές θέση στον τομέα αυτό, επιβαρύνοντας ακόμα περισσότερο την δημόσια εικόνα του. Τέλος, στη τέταρτη θέση βρίσκεται ο τομέας που περιέχει την μεταποίηση και την εξόρυξη.

Εικόνα 1: Οι 4 πιο διεφθαρμένοι τομείς σύμφωνα με την έρευνα



Η έρευνα δεν προσέφερε μόνο συμπεράσματα για το ποιοι τομείς έχουν τη περισσότερη διαφθορά. Συγκεκριμένα, εξάχθηκαν και ενδιαφέρουσες δηλώσεις σχετικά με τις άμεσες ξένες επενδύσεις, όπως για παράδειγμα το γεγονός ότι ο

τομέας της μεταποίησης και της εξόρυξης, παρότι δεν βρίσκεται στις πρώτες δύο θέσεις όσον αφορά τη διαφθορά, φαίνεται να παρουσιάζει ακόμα περισσότερα περιστατικά (διαφθοράς) όταν πρόκειται για ξένες-διεθνείς συμφωνίες.

5.2 Νέες μεταβλητές

Η Παγκόσμια Τράπεζα προσφέρει πολλά εργαλεία, ώστε κανείς να εξετάσει εάν συγκεκριμένοι τομείς με μεγάλη διαφθορά επηρεάζουν τα επίπεδα εισροών άμεσων ξένων επενδύσεων σε μία χώρα. Πιο συγκεκριμένα, θα χρησιμοποιηθούν δύο μέτρα. Το πρώτο αφορά τον δείκτη Industry Value ως ποσοστό του ΑΕΠ για κάθε χώρα. Η μέτρηση και ο όρος Industry μπορεί να ακούγεται κάπως γενικός αλλά είναι σε θέση να προσφέρει πολλές πληροφορίες έτσι ώστε να μπορεί να ερευνηθεί εάν ένα μεγάλο ποσοστό “διεφθαρμένων τομέων” επηρεάζει τις ροές FDI για τις συστάδες που εξετάζονται. Ο δείκτης “Industry value added” αφορά τα παραρτήματα ISIC 10-45 και περιέχει τον μεταποιητικό τομέα (ISIC divisions 15-37), τον εξορυκτικό, τον ενεργειακό και τον κατασκευαστικά. Όπως γίνεται κατανοητό αυτοί είναι και οι τομείς που αναφέρθηκαν ως διεφθαρμένοι στην έρευνα του οργανισμού και για αυτό επελέγη ο συγκεκριμένος δείκτης. Η μεταβλητή θα ονομαστεί IND.

Ως δεύτερο μέτρο, θα αφαιρεθεί ο μεταποιητικός τομέας από την παραπάνω μεταβλητή. Δηλαδή θα αφαιρεθούν τα παραρτήματα 15-37 από το σύνολο 10-45. Ο λόγος είναι πως ο μεταποιητικός τομέας είναι σχετικά κυρίαρχος στον παραπάνω δείκτη και επειδή σύμφωνα με τις απαντήσεις των συνεντεύξεων σε σχέση με τους άλλους τομείς του δείκτη είναι λιγότερο διεφθαρμένος, μπορεί αυτός να δημιουργήσει λανθασμένα συμπεράσματα. Η μεταβλητή θα ονομαστεί INDMM και θα περιέχει τα παραρτήματα 10-14 και 38-45. Ταυτόχρονα, θα αναλυθεί και ο μεταποιητικός τομέας ξεχωριστά και η μεταβλητή θα ονομαστεί MAN.

Εάν η διαφθορά επηρεάζει αρνητικά τις ροές FDI, θα περιμέναμε πως σε περίπτωση που αυτές οι διεφθαρμένες βιομηχανίες καταλαμβάνουν μεγάλο ποσοστό του ΑΕΠ μίας χώρα οι εισροές ξένων επενδύσεων να είναι μικρότερες. Το αντίθετο θα συνέβαινε σε περίπτωση θετικής σχέσης.

5.3 Μεθοδολογία

Τα δεδομένα θα χρησιμοποιηθούν σε ένα πλαίσιο παλινδρομήσεων και οι παραπάνω μεταβλητές θα εισαχθούν σε μοντέλα ως ανεξάρτητες, έχοντας ως εξαρτημένη φυσικά τις ροές FDI. Ως δεύτερο βήμα θα εισαχθεί και η μεταβλητή του δείκτη CPI στο μοντέλο. Αφού οι νέες μεταβλητές δεν αφορούν κάθε χώρα ξεχωριστά, θα είχε ιδιαίτερο ενδιαφέρον αν εξετάσουμε τη σχέση τους με τον δείκτη CPI, που αφορά country-level data.

Έτσι δημιουργείται το παρακάτω μοντέλο:

$$FDI_{it} = a + \beta x_{it} + \gamma z_{it} + \mu_i + e_{it} \quad (3)$$

Όπου μ_i συμβολίζει τον παράγοντα country fixed effect. Ο παράγων x συμβολίζει τον παραδοσιακό δείκτη διαφθοράς CPI. Υψηλότερες τιμές του x , αφορούν υψηλότερες τιμές για τη διαφθορά. Ομοίως, το z παριστά το ποσοστό του ΑΕΠ που προέρχεται από διεφθαρμένο τομέα, όπως αυτοί αποτυπώθηκαν στην έρευνα του οργανισμού Transparency International. Πιο συγκεκριμένα το z αφορά τις μεταβλητές IND, INDMM and MAN. Για να συνδεθεί το μοντέλο αυτό με τις υποθέσεις που αναφέρθηκαν παραπάνω καταλήγουμε στο παρακάτω:

Για $H2$: $\gamma < 0$ σημαίνει πως μεγαλύτερο ποσοστό του ΑΕΠ προερχόμενο από διεφθαρμένους τομείς οδηγεί σε μικρότερες ροές FDI.

Το μοντέλο αφορά δύο συστάδες 19 χωρών για μία περίοδο 18 ετών. Πρόκειται για panel model και χρησιμοποιούνται όπως και στα υπόλοιπα μοντέλα fixed effects.

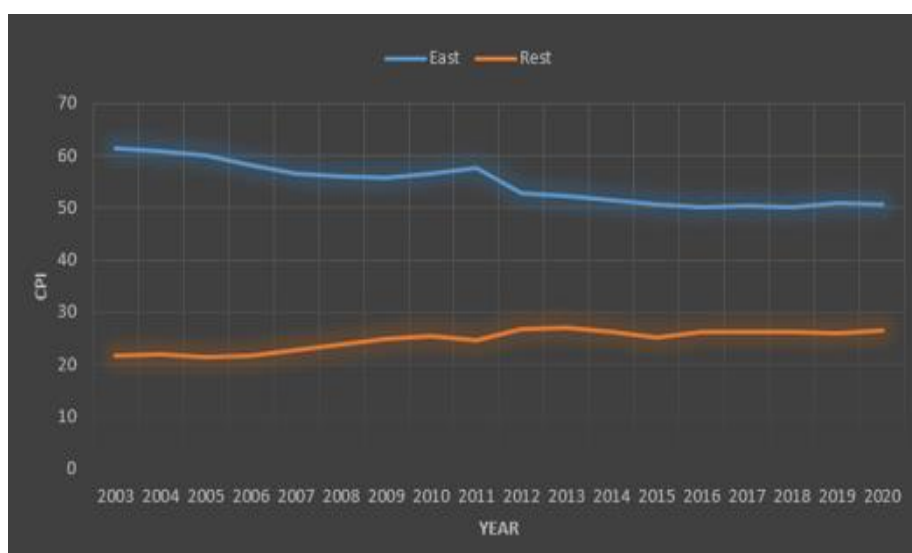
Αυτή η εναλλακτική προσέγγιση για την εξέταση της σχέσης μεταξύ FDI και διαφθοράς στοχεύει στην επιβεβαίωση ή την αμφισβήτηση των αποτελεσμάτων που βρέθηκαν στο πρώτο μέρος της ανάλυσής μου και παρέχει επίσης ένα διαφορετικό πλαίσιο για την εξέταση της διαφθοράς χωρίς τη χρήση δημοφιλών δεικτών όπως ο δείκτης CPI.

Η ερευνητική μου δομή βασίζεται στην ιδέα ότι η βιβλιογραφία δεν προσφέρει ένα συνεπές θεωρητικό πλαίσιο για την καθοδήγηση της εργασίας, έτσι ώστε να υπάρχει ενιαίο μοντέλο που να καθορίζει πλήρως τις μεταβλητές που μπορούν να θεωρηθούν σταθερές προκειμένου να διερευνηθούν οι επιπτώσεις της διαφθοράς, δηλαδή μεταβλητές που να συμπεριλαμβάνονται πάντα στα μοντέλα. Επομένως, πρέπει να εκτιμηθούν όλες οι πιθανές παλινδρομήσεις που μπορούν να καθοριστούν με την προσθήκη οποιουδήποτε συνδυασμού ανεξάρτητων μεταβλητών.

6. Αποτελέσματα Ανάλυσης Χρονοσειρών

Πριν την ανάλυση των μοντέλων παλινδρόμησης που δημιουργήθηκαν, σε αυτή την ενότητα θα εκπονηθεί η ανάλυση χρονοσειράς για τις δύο συστάδες χωρών και θα παρουσιαστεί μοντέλο πρόβλεψης ARIMA. Είναι σημαντικό να αναφέρουμε πως οι χρονοσειρές και τα μοντέλα πρόβλεψης αφορούν την εξαρτημένη μεταβλητή, δηλαδή τις εισροές FDI. Πριν την παρουσίαση των αποτελεσμάτων, θα είχε ενδιαφέρον να δούμε πως κινείται αντίστοιχα για τα έτη 2003-2020 ο δείκτης CPI για τις δύο συστάδες.

Γράφημα 3: Ετήσιοι μέσοι όροι των δύο συστάδων για τον δείκτη CPI για τα έτη 2003-2020.



Στο παραπάνω γράφημα οι γραμμές παρουσιάζουν την πορεία του μέσου όρου για τον δείκτη CPI για κάθε συστάδα. Η μπλε γραμμή αφορά την συστάδα που αποτελείται από τις χώρες της ανατολικής Ευρώπης και η πορτοκαλί την συστάδα του

υπολοίπου Ευρώπης. Αυτό που παρατηρούμε είναι πως τις τελευταίες δύο δεκαετίες το εύρος της διαφοράς όσον αφορά τη διαφθορά ανάμεσα στις δύο συστάδες φαίνεται να μειώνεται. Αυτό ως πρώτο συμπέρασμα είναι κάτι αρκετά θετικό, καθώς πολλές από τις χώρες της ανατολικής Ευρώπης βρίσκονταν ιδιαίτερα τις προηγούμενες δεκαετίες σε ένα μετα-κομμουνιστικό, μεταβατικό στάδιο και σαν πρώτη ανάγνωση φαίνεται πως οι νέοι νόμοι και οι κανόνες αυτών των χωρών έχουν επηρεάσει ως ένα βαθμό τα επίπεδα διαφθοράς. Κάτι που μπορεί να έχει παίξει ρόλο είναι φυσικά αυτό που είναι γνωστό ως Europeanisation και αφορά κυρίως τις χώρες της άλλης συστάδας, αλλά δεν πρέπει να ξεχνάμε πως πολλές χώρες των δύο συστάδων συνορεύουν μεταξύ τους και έτσι δεν πρέπει να αποκλείονται πιθανά spillovers αυτού του φαινομένου και στις χώρες της ανατολικής Ευρώπης. Για το έτος 2003 το εύρος μεταξύ των δύο συστάδων για το δείκτη CPI ήταν 40 μονάδες. Πιο συγκεκριμένα οι χώρες της ανατολικής συστάδας είχαν κατά μέσο όρο 40 μονάδες υψηλότερη διαφθορά, σύμφωνα με τον δείκτη, σε σχέση με τις χώρες της άλλης συστάδας. Όμως παρατηρούμε ότι το έτος 2020 το εύρος ανάμεσα στις συστάδες έχει μειωθεί στις 25 μονάδες, κυρίως εξαιτίας της σταδιακής μείωσης της διαφθοράς στην ανατολική συστάδα αλλά και της μικρής αύξησης του δείκτη διαφθοράς στην χώρες του υπολοίπου Ευρώπης. Από την μείωση λοιπόν αυτή των 15 μονάδων στο εύρος, ένα γρήγορο συμπέρασμα, αν όχι επιφανειακό, θα ήταν πως η διαφθορά στην ανατολική Ευρώπη μειώνεται και η κατάσταση στους νόμους και γενικότερα στις δομές βελτιώνεται. Υπάρχουν παράγοντες όμως που πρέπει να ληφθούν υπόψιν πριν τα τελικά συμπεράσματα. Ο πρώτος είναι πως η μείωση του εύρους σταματά στο έτος 2012. Άρα το σύνολο αυτής της μείωσης αφορά τα έτη 2003-2012 και όχι το σύνολο των ετών. Τα τελευταία 8 έτη της έρευνας, συγκεκριμένα τα έτη 2012-2020, το εύρος ανάμεσα στις δύο συστάδες παραμένει σταθερό, όπως φαίνεται και από το γράφημα. Άρα εάν υποθέσουμε πως η μείωση αυτή οφείλεται στην γενικότερη βελτίωση των ανατολικών χωρών, τότε αυτή σταμάτησε το 2012. Η πραγματικότητα όμως είναι πως αυτή η μείωση στον μέσο όρο που παρατηρείται δεν οφείλεται στο σύνολο των χωρών, αλλά στις τεράστιες αλλαγές που έγιναν σε μέρος της ανατολικής συστάδας κατά την δεκαετία 2000-2010. Πιο συγκεκριμένα οι χώρες που περιόρισαν σε τεράστιο βαθμό την διαφθορά και άρα μείωσαν τον δείκτη CPI, είναι αυτές της Βαλτικής, η Ρουμανία και η Πολωνία. Η βελτίωση στις χώρες της Βαλτικής οφείλεται

σε μεγάλο βαθμό σε μία έννοια που αναφέρθηκε παραπάνω, πιο συγκεκριμένα το φαινόμενο του Europeanisation. Οι χώρες αυτές είναι εκείνες με το μεγαλύτερο βαθμό Europeanisation από το σύνολο των χωρών της ανατολικής συστάδας. Η Εσθονία, η Λιθουανία και η Λετονία κατά τις πρόσφατες δεκαετίες άλλαξαν κατά πολύ μεγάλο βαθμό τις δομές κράτους και μάλιστα κάποιες από αυτές, όπως η Εσθονία αποτέλεσε πρότυπο σε αρκετούς τομείς ακόμα και για χώρες της δυτικής συστάδας. Το φαινόμενο του Europeanisation ίσως να πήρε αυτήν την έκταση σε αυτή τη γεωγραφική περιοχή λόγω των κακών σχέσεων αυτών των χωρών με τη Ρωσία και φυσικά την εγγύτητα τους με αυτήν.

Όσον αφορά την Πολωνία, πρόκειται για μία χώρα που και αυτή έχει αλλάξει πολλά πράγματα στις δομές τις, με μία τεράστια αύξηση του ΑΕΠ, κοντά στο 380%, τις τελευταίες δύο δεκαετίες. Μάλιστα δεν είναι λίγοι αυτοί που ανέφεραν πως η Πολωνία στο μέλλον της Ευρώπης μπορεί να αποκτήσει αντίστοιχο ρόλο με αυτόν της Γερμανίας, λόγω των πολλών πλεονεκτημάτων της, όπως το εργατικό δυναμικό και η τεχνολογία. Γίνεται ιδιαίτερα κατανοητό πως η Πολωνία λόγω και της γεωγραφικής της θέσης, κοντά δηλαδή στην κεντρική Ευρώπη, έχει πλέον πάρει διαφορετική πορεία σε σχέση με τις περισσότερες χώρες της ανατολικής συστάδας.

Τέλος, η Ρουμανία, μία χώρα με άλλοτε πολύ υψηλά επίπεδα διαφθοράς, έχει καταφέρει να την περιορίσει, εκτινάσσοντας επίσης και το δικό της ΑΕΠ. Πολύ σημαντικό ρόλο στις αλλαγές των δομών της Πολωνίας και της Ρουμανίας έπαιξε, όπως και στις χώρες της Βαλτικής, το φαινόμενο του Europeanisation, καθώς και οι δύο χώρες κατέβαλλαν προσπάθειες και τελικά κατάφεραν να μπουν στην Ευρωπαϊκή Ένωση στα μέσα της δεκαετίας 2000-2010.

Καταλήγοντας, το μεγαλύτερο μέρος της μείωσης των 40 μονάδων κατά τη διάρκεια της περιόδου 2003-2020 οφείλεται στα πρώτα χρόνια, μέχρι το 2012 και κατά πολύ μεγάλο βαθμό στις αλλαγές που συνέβησαν σε περιορισμένο αριθμό χωρών, συγκεκριμένα στο 25% της συστάδας και σε καμία περίπτωση δεν άφορα το σύνολο των χωρών.

Όπως όλη η ανάλυση αυτής της διπλωματικής εργασίας έτσι και αυτή η ενότητα θα εκπονηθεί μέσω της γλώσσας R και των εργαλείων που αυτή παρέχει. Η γλώσσα R

θεωρείται ιδανική για ανάλυση χρονοσειρών. Στην παρακάτω εικόνα δίνεται ένα δείγμα των δεδομένων, έτσι όπως διαμορφώθηκαν για την κατασκευή των χρονοσειρών. Κύρια πηγή, όπως έχει αναφερθεί ήδη για τις εισροές άμεσων ξένων επενδύσεων είναι η βάση δεδομένων της παγκόσμιας τράπεζας.

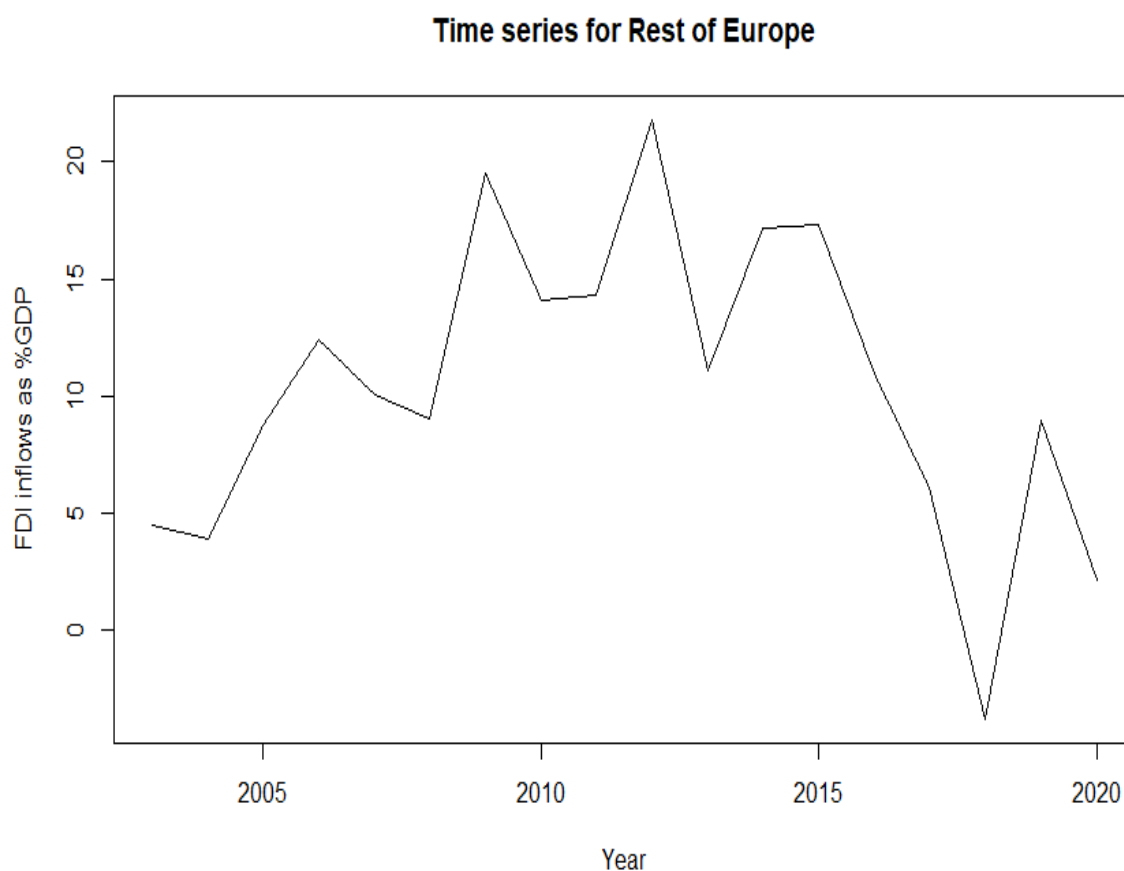
Εικόνα 2: Δείγμα δεδομένων στην R

<u>Rest of Europe</u>			<u>Eastern Europe</u>		
▲	Year	AVERAGE FDI	▲	Year	AVERAGE FDI
1	2003	4.500	1	2003	4.77777
2	2004	3.880	2	2004	5.73900
3	2005	8.720	3	2005	8.37500
4	2006	12.430	4	2006	7.35300
5	2007	10.024	5	2007	11.09800
6	2008	9.001	6	2008	8.76400
7	2009	19.520	7	2009	4.73500
8	2010	14.110	8	2010	3.44800
9	2011	14.330	9	2011	4.66900
10	2012	21.770	10	2012	3.73200
11	2013	11.054	11	2013	2.75400
12	2014	17.150	12	2014	3.55590
13	2015	17.280	13	2015	3.02800
14	2016	10.910	14	2016	6.12900
15	2017	6.094	15	2017	3.28900
16	2018	-3.770	16	2018	1.36200
17	2019	8.940	17	2019	3.55100
18	2020	2.147	18	2020	3.88100

Ως πρώτο βήμα επιλέγεται η κατασκευή της χρονοσειράς που αφορά την συστάδα του υπολοίπου Ευρώπης. Αφορά φυσικά τα χρόνια 2003-2020, όπως και το γράφημα που παρουσιάστηκε για τον δείκτη CPI. Για τις χώρες αυτές παρατηρούμε πως οι εισροές FDI ταυτίζονται σε ένα βαθμό με τις παγκόσμιες οικονομικές εξελίξεις των

τελευταίων δύο δεκαετιών. Η πρώτη μείωση-πτώση παρατηρείται κατά τη διάρκεια του έτους 2007, οπότε ξεκίνησε η παγκόσμια οικονομική κρίση. Τα επόμενα χρόνια, οι εισροές άμεσων ξένων επενδύσεων στην δυτική Ευρώπη ανέκαμψαν, ενώ φαίνεται το έτος 2012 να φτάνει και ιστορικό υψηλό. Μετά το 2016 ακολουθεί μία ραγδαία πτώση στις εισροές, και αυτό ίσως να έχει να κάνει με το γεγονός πως πολλές από τις χώρες της δυτικής Ευρώπης, έχουν ως κύριο ξένο επενδυτή τις Ηνωμένες Πολιτείες. Από το 2016 οι Η.Π.Α ακολούθησαν μία διαφορετική πολιτική όσον αφορά τις επενδύσεις σε ξένες χώρες και αυτή ίσως να είναι η αιτία για την μείωση των ροών. Τέλος, αυτό που προκαλεί ενδιαφέρον, είναι πως μετά το ιστορικό χαμηλό του 2018, οι ροές φάνηκαν να ανακάμπτουν ως ένα βαθμό τον επόμενο χρόνο, όμως η τάση αυτή δεν κράτησε πολύ, καθώς και τον χρόνο που ακολούθησε υπήρξε μείωση. Αυτή η τελευταία πτώση, είναι αρκετά πιθανό να συνδέεται και με τις επιδράσεις του κορονοϊού, αφού λόγω κλειστών συνόρων και περιορισμών πολλές επενδύσεις καθυστέρησαν ή δεν ολοκληρώθηκαν.

Γράφημα 4: Χρονοσειρά υπολοίπου Ευρώπης



Από το γράφημα της χρονοσειράς, δεν παρατηρείται κάποια τάση, ανοδική ή καθοδική αλλά ούτε και σημάδια εποχικότητας-κυκλικότητας. Σε περίπτωση που συνέβαινε αυτό θα μπορούσαμε με τα εργαλεία που παρέχει η γλώσσα R, να αφαιρέσουμε τέτοιου είδους επιδράσεις.

Εκτός όμως από την ερμηνεία του γραφήματος, εφαρμόστηκε και ο έλεγχος Dickey-Fuller για την στασιμότητα. Τα αποτελέσματα του οποίου παρουσιάζονται στον παρακάτω πίνακα.

Εικόνα 3: DF test για υπόλοιπο Ευρώπης

```
Augmented Dickey-Fuller Test
data: data
Dickey-Fuller = -3.3405, Lag order = 3, p-value = 0.08604
alternative hypothesis: stationary
```

Η μηδενική υπόθεση σύμφωνα με το p-value του πίνακα απορρίπτεται, πράγμα που σημαίνει ότι η χρονοσειρά είναι στάσιμη. Ο έλεγχος λοιπόν επιβεβαιώνει την αρχική ερμηνεία της χρονοσειράς με βάση το σχήμα.

Θέλοντας να επιβεβαιώσουμε τα αποτελέσματα θα χρησιμοποιηθεί και ο έλεγχος KPSS. Τα αποτελέσματα του οποίου δίνονται παρακάτω:

Εικόνα 4: KPSS test για υπόλοιπο Ευρώπης

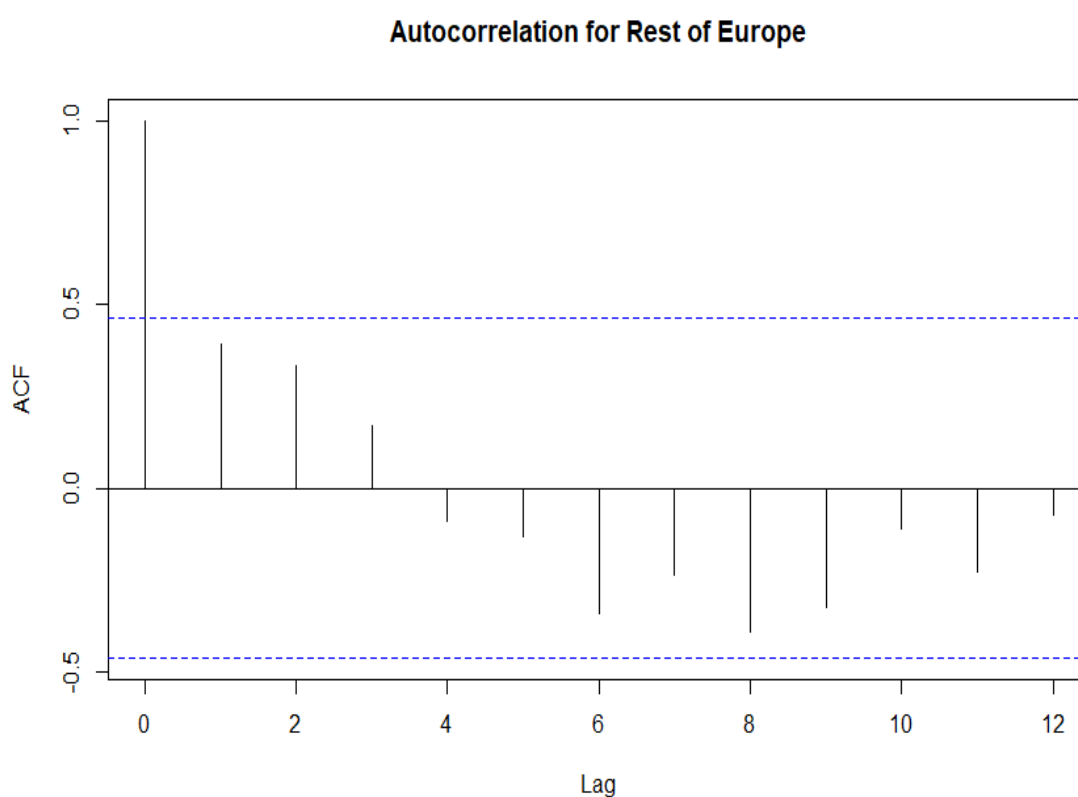
```
KPSS Test for Level Stationarity
data: data
KPSS Level = 0.30937, Truncation lag parameter = 2, p-value = 0.1
```

Για τον έλεγχο αυτόν, σε περίπτωση απόρριψης της μηδενικής υπόθεσης, η χρονοσειρά είναι μη στάσιμη. Παρατηρούμε πως το p-value είναι στο 0.1 άρα η

μηδενική υπόθεση για τα δεδομένα αυτά δεν μπορεί να απορριφθεί για κανένα επίπεδο σημαντικότητας α , το οποίο θα μπορούσε να ληφθεί υπόψιν. Καταλήγουμε λοιπόν και πάλι πως η χρονοσειρά που αφορά την συστάδα του υπολοίπου είναι στάσιμη.

Στο σημείο αυτό θα είχε ενδιαφέρον μέσω των εργαλείων της R, να ελεγχθεί και η συνάρτηση αυτοσυσχέτισης, καθώς μέσα από αυτό μπορούμε να λάβουμε αρκετές πληροφορίες και για την στασιμότητα.

Γράφημα 5: Αυτοσυσχετίσεις υπολοίπου Ευρώπης.



Αυτό που γίνεται σαφές από το γράφημα αυτοσυσχετίσεων, είναι ότι δεν υπάρχει καμία σημαντική συσχέτιση, πέρα φυσικά από την παρατήρηση, με τον εαυτό της. Όσον αφορά τις υπόλοιπες κανένας συντελεστής αυτοσυσχέτισης δεν ξεφεύγει πάνω από το 0.4. Τα αποτελέσματα αυτά είναι ακόλουθα φυσικά με το γεγονός πως η χρονοσειρά για τις χώρες του υπολοίπου Ευρώπης είναι στάσιμη. Θα περιμέναμε μέτριες και ισχυρές κυρίως συσχετίσεις στην περίπτωση μη στάσιμης χρονοσειράς.

Τα παραπάνω μας επιτρέπουν να προχωρήσουμε στην ανάλυση της χρονοσειράς μέσω μοντέλων ARIMA, έχοντας ως στόχο την δημιουργία προβλέψεων. Μέσω της R,

μπορούμε με πολύ εύκολο και άμεσο τρόπο, αντί να προβαίνομε σε αλληπάλληλες δοκιμές, να δημιουργήσουμε πίνακα, ο οποίος θα δηλώνει ποιο μοντέλο ARIMA και ποιοι συντελεστές είναι οι καταλληλότεροι. Η επιλογή αυτή θα γίνει με βάση το κριτήριο AIC – Akaike Information Criteria και στόχος είναι η επιλογή του μοντέλου με το μικρότερο συντελεστή για το AIC. Στον παρακάτω πίνακα παρουσιάζεται το σύνολο των μοντέλων ARIMA, από τα οποία θα γίνει η τελική επιλογή το μοντέλο που θα εφαρμοστεί. Είναι φανερό πως το μοντέλο με τον μικρότερο συντελεστή για το

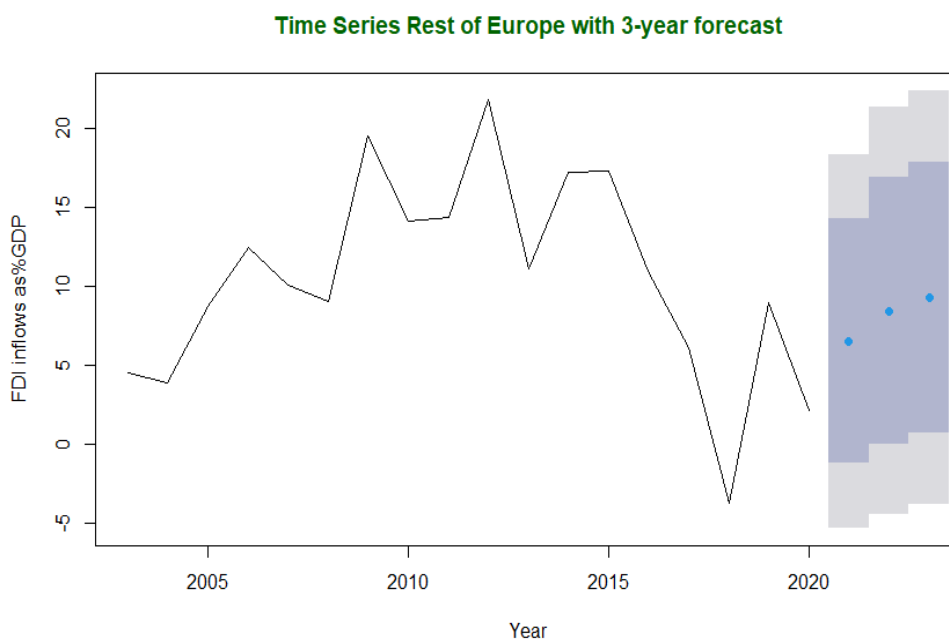
Εικόνα 5: AIC υπολοίπου Ευρώπης

```
ARIMA(2,0,2) with non-zero mean : Inf
ARIMA(0,0,0) with non-zero mean : 122.0587
ARIMA(1,0,0) with non-zero mean : 121.5601
ARIMA(0,0,1) with non-zero mean : 122.7637
ARIMA(0,0,0) with zero mean      : 143.3664
ARIMA(2,0,0) with non-zero mean : 123.8168
ARIMA(1,0,1) with non-zero mean : 124.117
ARIMA(2,0,1) with non-zero mean : 127.709
ARIMA(1,0,0) with zero mean     : 123.2461
```

```
Best model: ARIMA(1,0,0) with non-zero mean
```

κριτήριο AIC είναι το μοντέλο ARIMA(1,0,0). Είναι εύκολα κατανοητό ότι αυτό θα είναι και το μοντέλο ARIMA που θα εφαρμοστεί στην συνέχεια. Τα αποτελέσματα του μοντέλου ARIMA(1,0,0) απεικονίζονται στο παρακάτω γράφημα.

Γράφημα 6: ARIMA 3 ετών για υπόλοιπο Ευρώπης



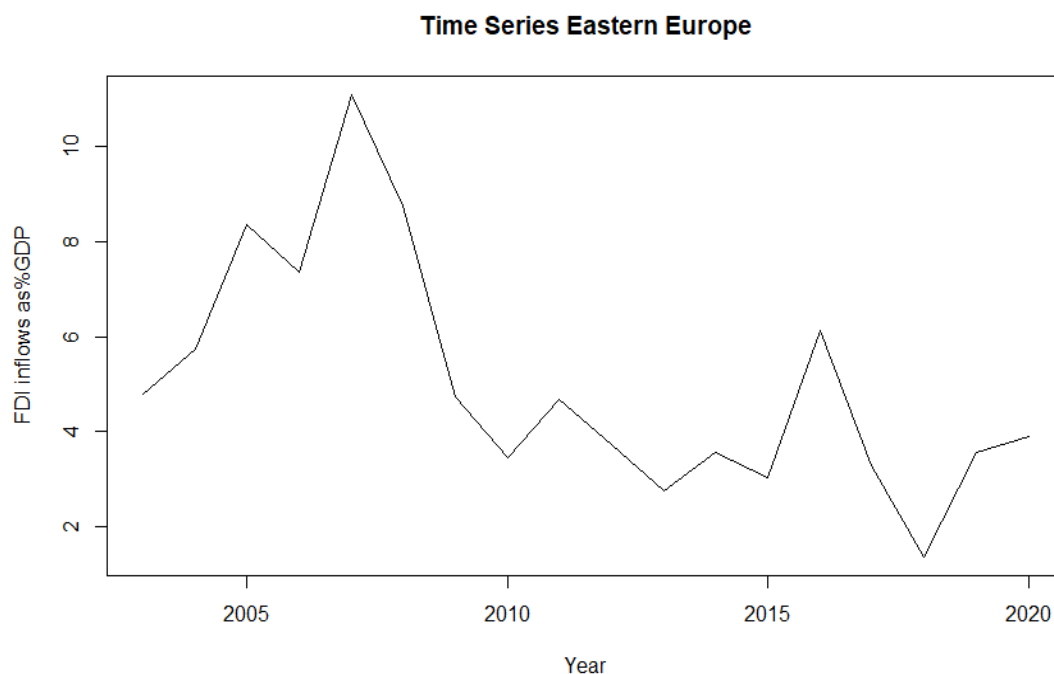
Παρά την συνολική πτωτική πορεία μετά το 2012 και ειδικά μετά την τελευταία πτώση του 2019, το μοντέλο ARIMA(1,0,0), προβλέπει πως οι εισροές άμεσων ξένων επενδύσεων ως ποσοστό του ΑΕΠ για τις χώρες του υπολοίπου Ευρώπης θα παρουσιάσουν αύξηση τα επόμενα τρία χρόνια. Πιο συγκεκριμένα η αύξηση αυτή φαίνεται να υπερκαλύπτει την πτώση του 2019. Το μέσο ποσοστό φτάνει το 9,23% για το έτος 2023 για την συγκεκριμένη συστάδα. Άρα από το 2,15% του έτους 2020 το μοντέλο προβλέπει αύξηση 7,08 ποσοστιαίων μονάδων. Η σημασία αυτής της πρόβλεψης, και συγκεκριμένα της αύξησης, γίνεται ακόμα πιο κατανοητή εάν σκεφτεί κανείς τι σημαίνει αυτό το αποτέλεσμα εάν μεταφραστεί σε δολάρια.

Αυτό που πρέπει να επισημανθεί πριν την παρουσίαση των αποτελεσμάτων για την συστάδα της ανατολικής Ευρώπης είναι πως η επιλογή πρόβλεψης τριών χρόνων μπροστά, θα μπορούσε μεν να περιγραφεί από κάποιον ως συντηρητική, όμως δεν πρέπει να ξεχνάμε πως τα μοντέλα ARIMA λόγω της δομής τους παρέχουν ακριβέστερες εκτιμήσεις στις βραχυχρόνιες προβλέψεις. Έτσι τα τρία χρόνια επελέγησαν ως μία βραχυπρόθεσμη εκτίμηση, που όμως μπορεί να παρέχει πολλές πληροφορίες.

Όσον αφορά τα αντίστοιχα αποτελέσματα για την συστάδα της Ανατολικής Ευρώπης παρατηρούμε από το γράφημα της χρονοσειράς που παρουσιάζεται παρακάτω πως

υπάρχει σαφής πτώση, η οποία φαίνεται να ξεκινάει το έτος 2007 όπου ξεκίνησε η παγκόσμια κρίση

Γράφημα 7: Χρονοσειρά ανατολικής Ευρώπης



Αυτή παρατηρήθηκε και στην συστάδα του υπολοίπου Ευρώπης, όμως στην περίπτωση των ανατολικών χωρών αυτή η μείωση φαίνεται να έχει πολύ μεγαλύτερη διάρκεια και για την ακρίβεια σχεδόν 8 χρόνια. Αυτή η ραγδαία πτώση φαίνεται να σταματάει το έτος 2015, όπου παρατηρείται μικρή ποσοστιαία αύξηση, την οποία όμως την διαδέχεται μία ακόμη πτώση. Μια διαφορά είναι πως οι επιδράσεις του κορονοϊού εδώ δεν φαίνεται να προκαλούν κάποια μείωση σε αντίθεση με την συστάδα του υπολοίπου, όπου το έτος που ξεκίνησε αυτό το πρόβλημα παρατηρήθηκε μείωση των ξένων επενδύσεων ως προς το ΑΕΠ. Όπως στην συστάδα της δυτικής Ευρώπης, έτσι και εδώ δεν παρατηρούνται επιδράσεις τάσης ή εποχικότητας, φαίνεται λοιπόν η χρονοσειρά να είναι και αυτή στάσιμη.

Για να επιβεβαιωθεί αυτό το αρχικό συμπέρασμα, όπως και πριν, εφαρμόστηκαν στατιστικοί έλεγχοι για την στασιμότητα. Τα αποτελέσματα του Dickey-Fuller test δίνονται παρακάτω.

Εικόνα 6: DF test για ανατολική Ευρώπη

```
Augmented Dickey-Fuller Test
data: data
Dickey-Fuller = -3.5183, Lag order = 3, p-value = 0.06135
alternative hypothesis: stationary
```

Σύμφωνα με τον έλεγχο και το p-value απορρίπτουμε την μηδενική υπόθεση και καταλήγουμε πως σύμφωνα και με αυτόν έλεγχο τη χρονοσειρά είναι στάσιμη.

Εικόνα 7: KPSS test για ανατολική Ευρώπη

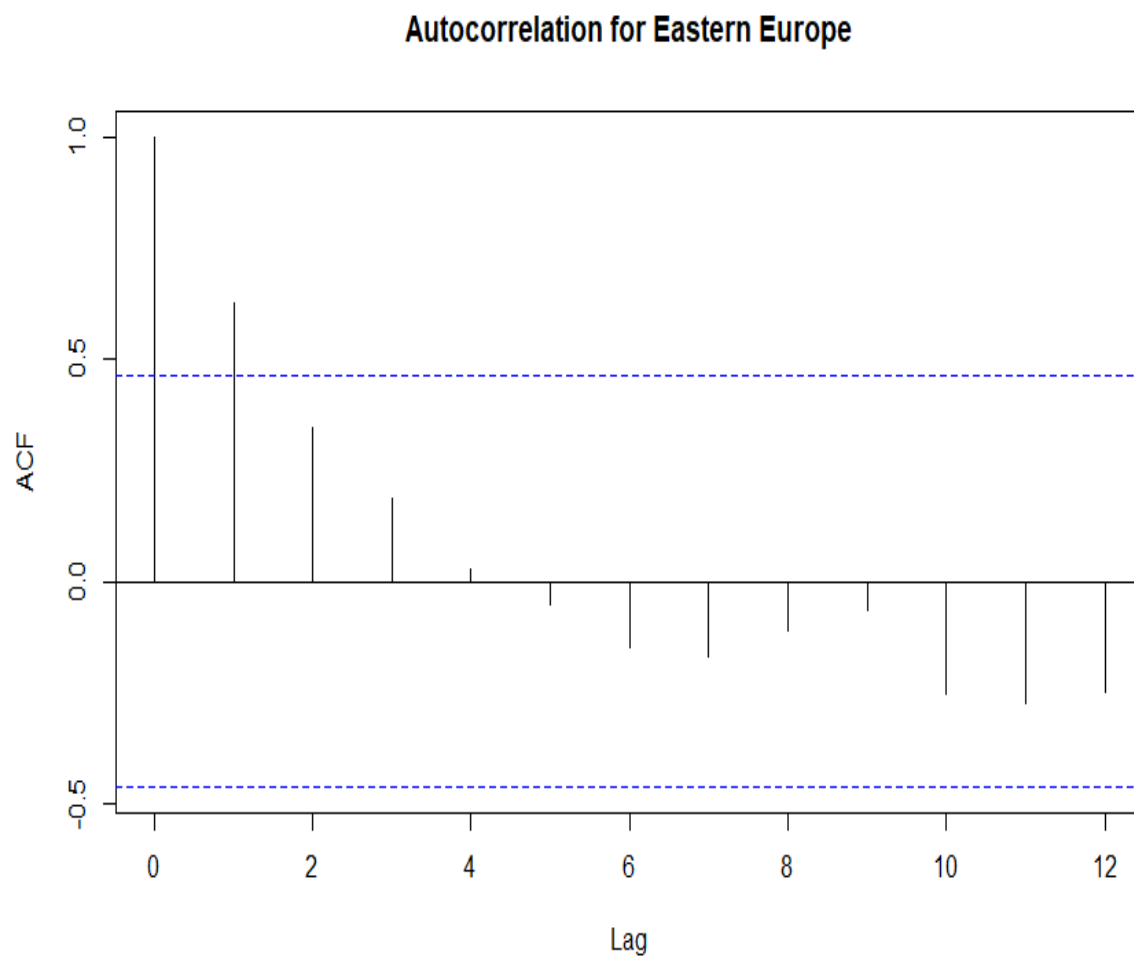
```
KPSS Test for Level Stationarity
data: data
KPSS Level = 0.10957, Truncation lag parameter = 2, p-value = 0.1
```

Με ανάλογο τρόπο εκπονήθηκε το KPSS test, όπου σύμφωνα με τα αποτελέσματα που δίνονται στον παραπάνω πίνακα δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση για κανέναν επίπεδο σημαντικότητας που θα γινόταν αποδεκτό και σύμφωνα με τις υποθέσεις του, καταλήγουμε πως η χρονοσειρά είναι στάσιμη.

Στο παρακάτω γράφημα παρουσιάζονται όπως και για την πρώτη συστάδα οι αυτοσυσχετίσεις. Όπως είναι λογικό, αφού πρόκειται για στάσιμη χρονοσειρά, δεν υπάρχουν ισχυρές συσχετίσεις. Η μόνη διαφορά με την άλλη συστάδα είναι πως υπάρχει μέτρια συσχέτιση με την ακριβώς επόμενη παρατήρηση, της τάξεως του 0,6. Όλες οι υπόλοιπες συσχετίσεις βρίσκονται κάτω από το 0,4. Τα αποτελέσματα αυτά είναι και τα αναμενόμενα, μετά και τα αποτελέσματα των ελέγχων που προηγήθηκαν.

Είναι σαφές και για τις δύο χρονοσειρές πως δεν υπάρχουν τέτοιες αυτοσυσχετίσεις, που θα μπορούσαν να θέσουν υπό αμφισβήτηση, ακόμα και χωρίς την παρουσία στατιστικών ελέγχων, το ότι οι χρονοσειρές είναι στάσιμες.

Γράφημα 8: Αυτοσυσχετίσεις ανατολικής Ευρώπης



Μετά τον έλεγχο για τη στασιμότητα μπορεί να ξεκινήσει η προβλεπτική διαδικασία

Στον παρακάτω πίνακα δίνονται και πάλι τα μοντέλα ARIMA με διάφορους συντελεστές, τα οποία αξιολόγησε η R. Σύμφωνα με το κριτήριο του AIC στα δεξιά του πίνακα το καταλληλότερο μοντέλο για αυτή την χρονοσειρά θα ήταν το ARIMA(0,1,0).

Εικόνα 8: AIC ανατολικής Ευρώπης

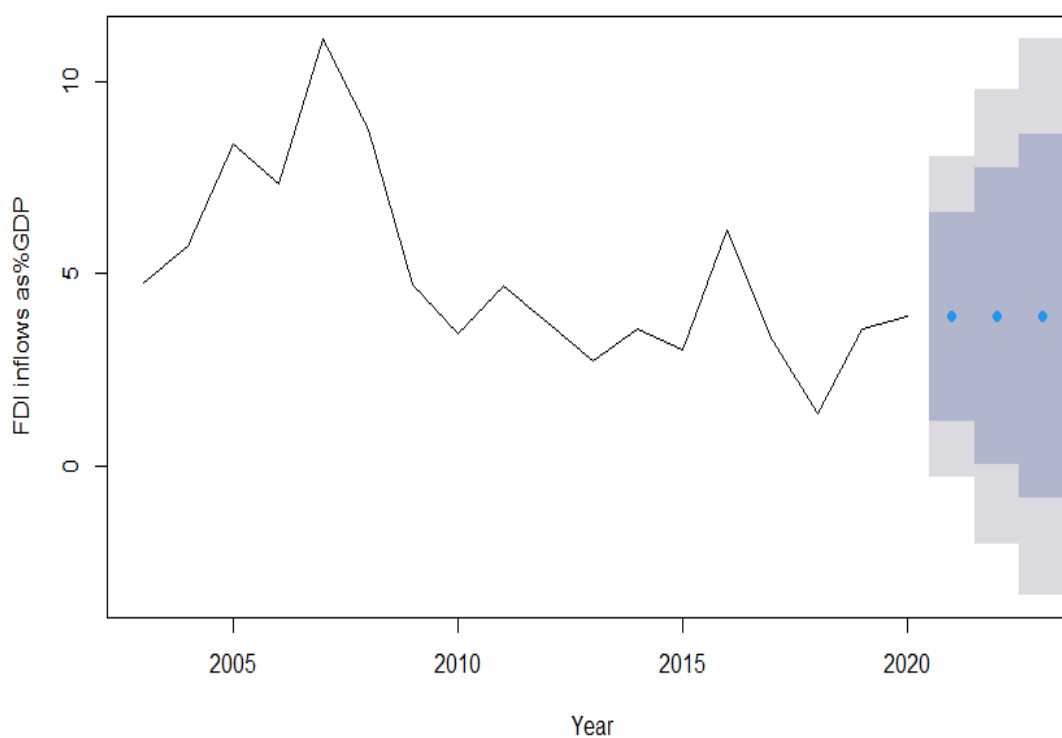
ARIMA(2,1,2) with drift	: Inf
ARIMA(0,1,0) with drift	: 78.73119
ARIMA(1,1,0) with drift	: 81.45592
ARIMA(0,1,1) with drift	: 81.24973
ARIMA(0,1,0)	: 76.15119
ARIMA(1,1,1) with drift	: Inf

Best model: ARIMA(0,1,0)

Το μοντέλο που επιλέχθηκε με βάση την παραπάνω αξιολόγηση είναι φυσικά και αυτό που εφαρμόστηκε. Τα αποτελέσματα του μοντέλου ARIMA(0,1,0) για την περιγραφή της χρονοσειράς και την πρόβλεψη των τριών χρόνων δίνεται στο επόμενο γράφημα.

Γράφημα 9: ARIMA 3 ετών ανατολικής Ευρώπης

Time Series Eastern Europe with 3-year forecast



Τα επόμενα τρία έτη, όπως προβλέπει το μοντέλο, δεν θα προκύψει κάποια ποσοστιαία αύξηση στις άμεσες ξένες επενδύσεις για τις χώρες της ανατολικής Ευρώπης συνολικά. Παρατηρούμε ότι για τα έτη 2021-2023 το μοντέλο προβλέπει οι εισροές FDI να παραμένουν σταθερές. Αυτή σαφώς είναι μία δυσάρεστη πρόβλεψη καθώς όπως έχει τονιστεί ήδη η περιοχή αυτή βρίσκεται σε τεράστια ανάγκη ξένων άμεσων επενδύσεων με στόχο την αύξηση των ΑΕΠ και του πλούτου των πολιτών της.

Εάν συγκρίνουμε τις ροές FDI με βάση τις χρονοσειρές και τα προβλεπτικά μοντέλα αλλά και τα γραφήματα που κατασκευάστηκαν με βάση τον δείκτη CPI, παρατηρούμε πως για την ανατολική Ευρώπη η μείωση της διαφθοράς τα τελευταία 20 χρόνια δεν οδήγησε σε καμία περίπτωση σε μεγαλύτερες ροές FDI ως ποσοστό του ΑΕΠ. Εάν η διαφθορά είχε αρνητική επίδραση στις ξένες επενδύσεις θα περιμέναμε πως με την μείωση αυτή, οι χώρες της ανατολικής Ευρώπης θα παρουσίαζαν μεγαλύτερα ποσοστά ξένων επενδύσεων ως ποσοστό του ΑΕΠ. Ένα αρκετά επιφανειακό συμπέρασμα θα ήταν πως αυτό είναι ένδειξη για το ότι η διαφθορά δεν έχει αρνητική επίδραση στις επενδύσεις. Όμως πιθανότατα τα αποτελέσματα των γραφημάτων επιβεβαιώνουν την άποψη πως η μείωση στην διαφθορά που παρατηρήθηκε δεν αφορά το σύνολο των ανατολικών χωρών, αλλά μάλλον δημιουργείται κυρίως λόγω της ραγδαίας μείωσης σε ένα μικρό μόνο μέρος της συστάδας.

Τα μοντέλα ARIMA που δημιουργήθηκαν με βάση τις χρονοσειρές προβλέπουν την αύξηση των εισροών επενδύσεων στις χώρες του υπολοίπου Ευρώπης, ενώ για τις χώρες της ανατολικής συστάδας αυτές φαίνεται να μένουν σταθερές για τα επόμενα έτη. Δυστυχώς όμως, όσον αφορά τις χώρες της ανατολής φαίνεται να παραμένουν σταθερές σε πολύ χαμηλό ποσοστό ως προς το ΑΕΠ και σίγουρα μακριά από τους στόχους και τις ανάγκες των χωρών που αποτελούν αυτήν την συστάδα. Αυτές οι προβλέψεις αναδεικνύουν και αυτές με τη σειρά τους την ανάγκη που υπάρχει για να απαντηθεί ο γρίφος του αν η διαφθορά επηρεάζει τις επενδύσεις και γενικότερα ποιοι άλλοι παράγοντες είναι σημαντικοί στην λήψη αποφάσεων των επενδυτών για την κατανομή των πόρων τους.

7.Αποτελέσματα Παλινδρομήσεων

7.1 Συμβατικό Μοντέλο

Η πρώτη αναφορά σε αποτελέσματα δεν γίνεται μόνο για την εξαγωγή συμπερασμάτων αλλά και για να αναλυθεί οπτικά μία από τις υποθέσεις του γραμμικού μοντέλου. Πιο συγκεκριμένα, ο πίνακας 1 που παρουσιάζεται παρακάτω δίνει τα αποτελέσματα του απλού γραμμικού μοντέλου με εξαρτημένη μεταβλητή τις εισροές FDI και μοναδική ανεξάρτητη μεταβλητή τον δείκτη Cpi. Στο αριστερό μέρος του πίνακα δίνονται τα αποτελέσματα του μοντέλου που δημιουργήθηκε για το υπόλοιπο των Ευρωπαϊκών χωρών, ενώ στο δεξί μέρος παρουσιάζονται αντιστοίχως αυτά των χωρών της ανατολικής Ευρώπης. Μία από τις βασικές υποθέσεις του γραμμικού μοντέλου παλινδρόμησης είναι η ύπαρξη ομοσκεδαστικότητας στα δεδομένα. Στα παρακάτω αποτελέσματα όμως δεν έχουν ληφθεί υπόψιν οι επιδράσεις της ετεροσκεδαστικότητας σε περίπτωση που αυτή είναι παρούσα.

Πίνακας 1: Αρχικό Μοντέλο

Rest of Europe		Eastern Europe	
VARIABLES	(1) FDI	VARIABLES	(1) FDI
Cpi	-1.539*** (0.361)	Cpi	0.149* (0.062)
Fixed Effects	Yes	Fixed Effects	Yes
Observations	342	Observations	342
Number of Countries	19	Number of Countries	19
R-squared	0.054	R-squared	0.017
Adj R-squared	0.002	Adj R-squared	0.040
Notes: Fixed Effects are employed. Cpi refers to the CP index. Robust Standard errors in parentheses***significant at 0% level, **significant at 0.1% level, *significant at 1%, . significant at 5%		Notes: Fixed Effects are employed. Cpi refers to the CP index. Robust Standard errors in parentheses***significant at 0% level, **significant at 0.1% level, *significant at 1%, . significant at 5%	

Η επιλογή της παρουσίασης αυτού του πίνακα, ο οποίος περιέχει τις επιδράσεις της ετεροσκεδαστικότητας γίνεται για να γίνει κατανοητό το πόσο μπορεί να επηρεάσει αυτή τα αποτελέσματα και τα τελικά συμπεράσματα, σε περίπτωση που δεν ληφθεί υπόψιν. Αρχικά εάν μπορούμε στη διαδικασία της ερμηνείας αυτών των αποτελεσμάτων βλέπουμε ότι και για τις δύο συστάδες χωρών παίρνουμε στατιστικά σημαντικά αποτελέσματα για τον δείκτη της διαφθοράς. Πιο συγκεκριμένα, για την συστάδα χωρών του υπολοίπου Ευρώπης, λαμβάνουμε στατιστικά σημαντικά αποτελέσματα και απορρίπτουμε τη μηδενική υπόθεση σε επίπεδο σημαντικότητας 0%, ενώ σύμφωνα με τον συντελεστή της κλίσης με μία μείωση μίας μονάδας στο δείκτη Cpi οι εισροές άμεσων ξένων επενδύσεων αυξάνονται κατά 1.539 ποσοστιαίες μονάδες. Αντίστοιχα για τη συστάδα της ανατολικής Ευρώπης λαμβάνουμε θετικό coefficient – στατιστικά σημαντικό αποτέλεσμα σε επίπεδο 1%, που σημαίνει πως με μία αύξηση κατά μία μονάδα οι εισροές FDI θα αυξηθούν κατά 0.149 ποσοστιαίες μονάδες. Παρατηρούμε πως και για τα δύο μοντέλα έχουμε 342 παρατηρήσεις και 19 χώρες ενώ εφαρμόζονται Fixed Effects. Για την εφαρμογή των Fixed Effects διενεργήθηκαν οι κατάλληλοι έλεγχοι, δηλαδή F test, Hausman και Breusch-Pagan, έτσι ώστε να αποφασιστεί αν το καταλληλότερο μοντέλο κάθε φορά θα δοθεί μέσω OLS, Fixed Effects ή Random Effects. Κατα τη διάρκεια της ανάλυσης φάνηκε σύμφωνα με αποτελέσματα πως η επιλογή των Fixed Effects είναι η καταλληλότερη. Είναι ιδιαίτερα ενδιαφέρον πως παρατηρούμε αντίθετες επιδράσεις-πρόσημα τις διαφθοράς στις δύο συστάδες. Στις δυτικές χώρες φαίνεται να επιδρά αρνητικά η διαφθορά, ενώ στην ανατολική Ευρώπη η διαφθορά ελκύει κατά κάποιο τρόπο τους επενδυτές.

Σε αυτό το σημείο όμως είναι ώρα να αναφέρουμε πως παρότι τα αποτελέσματα είναι ιδιαίτερα ενδιαφέροντα και είναι σαφώς μία αρχή στην ανάλυση δεν είναι δυνατόν να εξαχθούν συμπεράσματα. Αυτό συμβαίνει για δύο κύριους λόγους. Ο πρώτος είναι φυσικά η έλλειψη μεταβλητών ελέγχου, δηλαδή να εξακριβωθεί το αν οι αυξομειώσεις στις άμεσες ξένες επενδύσεις οφείλονται στον δείκτη της διαφθοράς ή οφείλονται σε άλλους παράγοντες. Ο δεύτερος λόγος είναι φυσικά η ύπαρξη ετεροσκεδαστικότητας. Σε όλη τη διάρκεια της ανάλυσης έγινε έλεγχος για την

ετεροσκεδαστικότητα και σε πολλά μοντέλα απο αυτά που δημιουργήθηκαν ήταν παρούσα. Ως αποτέλεσμα μέσω των εργαλείων που παρέχει η R, μετά τον ανάλογο έλεγχο, δόθηκε η δυνατότητα να αφαιρεθεί από όλα τα μοντέλα η ετεροσκεδαστικότητα και οι επιδράσεις της. Ως εκ τούτου, σε όλα τα αποτελέσματα που θα ακολουθήσουν εφαρμόζονται Fixed Effects και αφαιρούνται οι επιδράσεις της ετεροσκεδαστικότητας, έχοντας ως στόχο όσο πιο ακριβή αποτελέσματα για την εξαγωγή συμπερασμάτων.

Πίνακας 2: Αρχικό Μοντέλο χωρίς ετεροσκεδαστικότητα

Rest of Europe		Eastern Europe	
VARIABLES	(1) FDI	VARIABLES	(1) FDI
Cpi	-1.539 (1.013)	Cpi	0.149 . (0.081)
Fixed Effects	Yes	Fixed Effects	Yes
Observations	342	Observations	342
Number of Countries	19	Number of Countries	19
R-squared	0.054	R-squared	0.017
Adj R-squared	0.002	Adj R-squared	0.040
Notes: Fixed Effects are employed. Cpi refers to the CP index. Robust Standard errors in parentheses***significant at 0% level, **significant at 0.1% level, *significant at 1%, . significant at 5%		Notes: Fixed Effects are employed. Cpi refers to the CP index. Robust Standard errors in parentheses***significant at 0% level, **significant at 0.1% level, *significant at 1%, . significant at 5%	

Στα αποτελέσματα του πίνακα 2 παρουσιάζεται το ίδιο μοντέλο, μόνο που αυτή τη φορά λείπουν οι επιδράσεις της ετεροσκεδαστικότητας. Φυσικά οι επιδράσεις του δείκτη στις εισροές άμεσων ξένων επενδύσεων είναι ίδιες όσον αφορά τα πρόσημα. Πιο συγκεκριμένα παρατηρούμε πάλι πως για το υπόλοιπο της Ευρώπης ο δείκτης Cpi παρουσιάζει αρνητικό coefficient, ενώ για τις χώρες της ανατολικής Ευρώπης ο αντίστοιχος συντελεστής είναι θετικός. Καθώς εφαρμόζονται και πάλι Fixed Effects, η κυριότερη διαφορά ανάμεσα στα αποτελέσματα του πίνακα 2, δηλαδή του πίνακα που αφαιρούνται οι επιδράσεις της ετεροσκεδαστικότητας και του πρώτου πίνακα

έχει να κάνει με το εάν τα αποτελέσματα είναι στατιστικά σημαντικά. Συγκρίνοντας με τον πίνακα 1 παρατηρούμε πως για το υπόλοιπο της Ευρώπης λαμβάνουμε στατιστικά μη σημαντικό αποτέλεσμα για τον δείκτη της διαφθοράς. Γίνεται άμεσα κατανοητό το πόσο σημαντικό είναι να λάβει κανείς υποψίν του κατα τη διάρκεια της ανάλυσης την επίδραση της ετεροσκεδαστικότητας στα αποτελέσματα. Στον πίνακα 1 όπου είναι ενσωματωμένες οι επιδράσεις το πιο βασικό συμπέρασμα για τις χώρες της πρώτης συστάδας θα ήταν πως η χαμηλή διαφθορά είναι εξαιρετικά σημαντική για την αύξηση των εισροών FDI. Μετά την αφαίρεση των επιδράσεων όμως το συμπέρασμα ενός αναλυτή αλλάζει εντελώς, καθώς δεν είναι στατιστικά σημαντικό για κανένα επίπεδο σημαντικότητας που θα μπορούσε να χρησιμοποιηθεί στην περίπτωση αυτή. Αντίθετα για τις χώρες της ανατολικής Ευρώπης παρατηρούμε πως ακόμα και μετά την αφαίρεση των επιδράσεων ο συντελεστής του δείκτη Cpi παραμένει στατιστικά σημαντικός, έστω και σε επίπεδο σημαντικότητας 5%. Όμως ακόμα και σε αυτήν την περίπτωση υπάρχει μεταβολή αφού στον πρώτο πίνακα είχαμε ένα αποτέλεσμα που μπορούσαμε να απορρίψουμε τη μηδενική υπόθεση σε επίπεδο σημαντικότητας 1%. Σύμφωνα με τον νέο πίνακα λοιπόν, σε επίπεδο σημαντικότητας 5%, μία αύξηση μίας μονάδας στον δείκτη διαφθοράς φέρνει αύξηση 0.149 ποσοστιαίας μονάδας στις εισροές άμεσων ξένων επενδύσεων.

Ένας απο τους στόχους αυτής της διπλωματικής εργασίας είναι να ερευνησει και τις γεωγραφικές διαφορές που μπορεί να προκύψουν στα αποτελέσματα, αφού μία απο τις αφορμές για την εκπόνηση της ήταν οι ιδιαιτερότητες της ανατολικής Ευρώπης και το οικονομικό και πολιτικό της παρελθόν. Εξαρχής πάντως φαίνεται να προκύπτουν κάποιες διαφορές, αφού στατιστικά σημαντικά αποτελέσματα για τον δείκτη της διαφθοράς και μάλιστα με θετική επίδραση στις επενδύσεις, προέκυψαν μόνο για την συστάδα της ανατολικής Ευρώπης.

Ακόμα και μετά την αφαίρεση όμως των επιδράσεων την ετεροσκεδαστικότητας απο το μοντέλο, δεν μπορούν να εξαχθούν τελικά συμπεράσματα. Πριν γίνει αυτό, τα αποτελέσματα πρέπει να ελεγχθούν, μέσω του εμπλουτισμού του μοντέλου με αριθμό μεταβλητών ελέγχου. Είναι σαφές και από τον συντελεστή προσαρμογής R^2 πως το μοντέλο είναι αρκετά ελλειπές για να ληφθούν αποφάσεις με βάση αυτό.

Στους επόμενους πίνακες εισάγονται οι κατάλληλες μεταβλητές ελέγχου στο μοντέλο. Στη παρούσα εργασία αυτό θα γίνει δημιουργώντας σύνολα μεταβλητών (vectors), όπως δηλαδή αναφέρθηκε και στην υποενότητα της μεθοδολογίας. Φυσικά για να γίνει αυτό έχει γίνει έλεγχος σχετικός με τις συσχετίσεις ανάμεσα στις μεταβλητές, καθώς υψηλή συσχέτιση μεταξύ των μεταβλητών μπορεί να οδηγήσει σε λάθος αποφάσεις απόρριψης και άρα λάθος συμπεράσματα.

Στον πίνακα 3α παρουσιάζονται τα αποτελέσματα για τα δύο μοντέλα πολλαπλής παλινδρόμησης που δημιουργήθηκαν με βάση τις μεταβλητές ελέγχου που σχετίζονται με την πολιτική σταθερότητα των χωρών και το επίπεδο κοινωνικών ελευθεριών που αυτές προσφέρουν στους πολίτες τους. Ο πίνακας αυτός αφορά τις χώρες της συστάδας που αφορά το υπόλοιπο της Ευρώπης.

Πίνακας 3α: Εισαγωγή μεταβλητών ελέγχου σχετικών με την πολιτική σταθερότητα και τις κοινωνικές ελευθερίες για το υπόλοιπο της Ευρώπης

Rest of Europe		
VARIABLES	(1) FDI	(2) FDI
Cpi	-1.511 (0.928)	-1.513 (0.932)
ROL	40.969 . (24.792)	41.078. (24.876)
VOICE	-27.603 . (15.482)	-27.408. (15.149)
POLSTAB		-0.642 (5.903)
Fixed Effects	Yes	Yes
Observations	342	342
R-squared	0.082	0.082
Adj R-squared	0.022	0.019
Number of Countries	19	19

Notes: Fixed Effects are employed. Cpi refers to the CP index. ROL refers to the trust of members of the society on laws. VOICE refers to the Voice and accountability of citizens. POLSTAB refers to political stability and Absence of violence. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Το πρώτο πράγμα που παρατηρεί κανείς στη στήλη 1 είναι πως μετά την εισαγωγή των δύο μεταβλητών ελέγχου η μεταβλητή που αντιπροσωπεύει τον δείκτη Cri, παρότι συνεχίζει να έχει παρόμοιο αρνητικό συντελεστή (coefficient), είναι στατιστικά μη σημαντική για όλα τα επίπεδα σημαντικότητας. Στη στήλη 1 επελέγησαν να εισαχθούν στο μοντέλο οι μεταβλητές ROL και VOICE. Η μεταβλητή ROL αντιπροσωπεύει τον δείκτη Rule of Law της παγκόσμιας τράπεζας και στην ουσία απεικονίζει το πόση εμπιστοσύνη έχουν οι πολίτες κάθε χώρας στους νόμους της. Σύμφωνα με τα αποτελέσματα απορρίπτουμε τη μηδενική υπόθεση σε επίπεδο εμπιστοσύνης 5% και καταλήγουμε πως η μεταβλητή αυτή είναι στατιστικά σημαντική για την εξαρτημένη μεταβλητή FDI. Πιο συγκεκριμένα, μία αύξηση 40.969 μονάδων στον δείκτη εμπιστοσύνης ROL οδηγεί σε αύξηση μίας ποσοστιαίας μονάδας για την μεταβλητή των εισροών άμεσων ξένων επενδύσεων.

Η δεύτερη μεταβλητή ελέγχου που εισάγεται στο μοντέλο, του οποίου τα αποτελέσματα απεικονίζονται στη στήλη 1 είναι η μεταβλητή VOICE. Αυτή απεικονίζει τις ελευθερίες των μέσων μαζικής ενημέρωσης, τις ελευθερίες των πολιτών σε θέματα όπως η συμμετοχή σε απεργίες και γενικότερα η ελευθερία έκφρασης. Εδώ λαμβάνουμε ένα σχετικά παράδοξο αποτέλεσμα, αφού ο συντελεστής είναι αρνητικός, δηλαδή η εξαρτημένη και η ανεξάρτητη αυτή μεταβλητή κινούνται προς την αντίθετη κατεύθυνση. Πιο συγκεκριμένα για τη μεταβλητή VOICE, απορρίπτουμε τη μηδενική υπόθεση και καταλήγουμε πως σε επίπεδο σημαντικότητας 5% είναι στατιστικά σημαντική ως προς το ύψος των άμεσων ξένων επενδύσεων. Για μία μείωση του δείκτη αυτού κατά μία μονάδα, οι εισροές FDI αυξάνονται κατά 27.603 ποσοστιαίες μονάδες. Μία υπόθεση που μπορεί να γίνει για αυτό το αποτέλεσμα είναι πως οι επενδυτές επικεντρώνονται πιο πολύ στην απόλυτη εφαρμογή των νόμων και των κανόνων, ενώ ίσως να είναι αρνητικά τα συναισθήματα τους για τις επιπτώσεις των απεργιών. Τα δύο τελευταία είναι σύμφωνα με το γεγονός ότι και η μεταβλητή ROL είναι στατιστικά σημαντική για τους ξένους επενδυτές.

Στη στήλη 2 του πίνακα 3α, το μοντέλο της στήλης 1 εμπλουτίζεται από την μεταβλητή ελέγχου POLSTAB. Η μεταβλητή αυτή απεικονίζει την βαθμολόγηση της παγκόσμιας

τράπεζας σχετικά με την πολιτική σταθερότητα κάθε χώρας. Το αποτέλεσμα αυτής της μεταβλητής δεν φαίνεται να είναι στατιστικά σημαντικό για το ύψος των εισροών ξένων επενδύσεων, αφού δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση για κανένα επίπεδο σημαντικότητας. Επιπρόσθετα, η εισαγωγή της μεταβλητής αυτής, δεν επηρεάζει σχεδόν καθόλου τις μεταβλητές ελέγχου ROL και VOICE. Πιο συγκεκριμένα και οι δύο παραμένουν στατιστικά σημαντικές για $\alpha = 5\%$, ενώ και οι συντελεστές (coefficients) παραμένουν σχεδόν ίδιοι, άρα τα συμπεράσματα που βγάλαμε στη στήλη 1 είναι τα ίδια με τη στήλη 2 για τις δύο μεταβλητές.

Στον πίνακα 3β παρουσιάζονται τα αποτελέσματα του εμπλουτισμένου μοντέλου για την συστάδα χωρών της ανατολικής Ευρώπης. Ο πίνακας 3β κινείται στο ίδιο μοτίβο με τον πίνακα 3α. Μία διαφορά με τον πίνακα 3α που αφορούσε τις χώρες του υπολοίπου της Ευρώπης, είναι στην αντιμετώπιση της μεταβλητής Cpi. Στις χώρες του Ευρωπαϊκού υπολοίπου αφού αφαιρέθηκαν οι επιδράσεις της ετεροσκεδαστικότητας, η μεταβλητή Cpi φάνηκε να είναι στατιστικά μη σημαντική. Αντίθετα για τις χώρες της ανατολικής Ευρώπης ακόμα και μετά την αφαίρεση της ετεροσκεδαστικότητας η επιδράσεις του δείκτη διαφθοράς στις άμεσες ξένες επενδύσεις παρέμειναν στατιστικά σημαντικές. Ως εκ τούτου, η ανάγκη για την εισαγωγή μεταβλητών ελέγχου στο μοντέλο που αφορά αυτή τη συστάδα φαίνεται να είναι πιο επιτακτική.

Στην στήλη 1 εισάγονται οι μεταβλητές ROL και VOICE. Το πρώτο πράγμα που παρατηρεί κανείς είναι πως πάρα τον εμπλουτισμό του μοντέλου με τις μεταβλητές ελέγχου η κύρια ανεξάρτητη μεταβλητή Cpi παραμένει στατιστικά σημαντική, όπως δηλαδή και στο αρχικό μοντέλο για τις χώρες της ανατολικής Ευρώπης. Πιο συγκεκριμένα, η μεταβλητή Cpi έχει θετικό συντελεστή, σε αντίθεση με τις χώρες του υπολοίπου της Ευρώπης. Αυτό το εύρημα, δηλαδή το γεγονός πως η εξαρτημένη μεταβλητή και η κύρια ανεξάρτητη μεταβλητή κινούνται προς την ίδια κατεύθυνση σημαίνει πως για τις χώρες αυτές υψηλά επίπεδα διαφθοράς φέρνουν μεγαλύτερες εισροές άμεσων ξένων επενδύσεων. Τέτοια ευρήματα είναι ιδιαίτερα συνηθισμένα στη βιβλιογραφία όπως ήδη αναφέρθηκε και στην ανάλογη υποενότητα. Οι διαφορές στα ευρήματα και στην αντιμετώπιση των επενδυτών ως προς τις επιλογές τους για την κατανομή των πόρων τους ήταν μία από τις αφορμές εκπόνησης της

διπλωματικής εργασίας αλλά και της επιλογής των συγκεκριμένων συστάδων χωρών. Για αυτό το μοντέλο, απορρίπτουμε τη μηδενική υπόθεση για $\alpha = 1\%$ που σημαίνει πως η μεταβλητή Cpi είναι στατιστικά σημαντική για την μεταβλητή FDI και καταλήγουμε πως με μία αύξηση του δείκτη κατά 0.193 μονάδες, οι εισροές FDI αυξάνονται κατά μία ποσοστιαία μονάδα.

Πίνακας 3β: Εισαγωγή μεταβλητών ελέγχου σχετικών με την πολιτική σταθερότητα και τις κοινωνικές ελευθερίες για την ανατολική Ευρώπη

Eastern Europe		
VARIABLES	(1) FDI	(2) FDI
Cpi	0.193* (0.090)	0.184* (0.083)
ROL	0.047 (2.442)	-1.139 (2.285)
VOICE	8.806 . (4.572)	9.058* (4.553)
POLSTAB		1.845* (0.872)
Fixed Effects	Yes	Yes
Observations	342	342
R-squared	0.050	0.056
Adj R-squared	0.011	0.009
Number of Countries	19	19

Notes: Fixed Effects are employed. Cpi refers to the CP index. ROL refers to the trust of members of the society on laws. VOICE refers to the Voice and accountability of citizens. POLSTAB refers to political stability and Absence of violence. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Στο μοντέλο αυτό της στήλης 1 η μεταβλητή VOICE, που αναφέρεται στις κοινωνικές ελευθερίες των πολιτών φαίνεται να είναι επίσης στατιστικά σημαντική σε επίπεδο σημαντικότητας 5% για την μεταβλητή FDI. Το γεγονός πως είναι στατιστικά

σημαντική αυτή η μεταβλητή παρατηρήθηκε και στις χώρες του υπολοίπου της Ευρώπης. Αυτό που προκαλεί μεγάλο ενδιαφέρον όμως είναι πως και εδώ, όπως και στην μεταβλητή Cri, παρατηρείται αντίθετος συντελεστής σε σχέση με την άλλη συστάδα. Πιο συγκεκριμένα μία αύξηση μίας μονάδας του δείκτη VOICE οδηγεί σε αύξηση 8.806 ποσοστιαίων μονάδων τις ξένες επενδύσεις. Όσον αφορά την μεταβλητή ROL, δεν φαίνεται να έχει κάποιο ρόλο στις αποφάσεις εκείνων που επενδύουν στις χώρες της ανατολικής Ευρώπης σε αντίθεση με εκείνους της πρώτης συστάδας.

Στη δεύτερη στήλη εισάγεται η μεταβλητή POLSTAB, που αναφέρεται στην πολιτική σταθερότητα και στα επίπεδα βίας στην κοινωνία. Η μεταβλητή αυτή διαφοροποιεί ακόμα περισσότερο τα αποτελέσματα των δύο συστάδων, αφού για τις χώρες της ανατολικής Ευρώπης φαίνεται να είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%. Πιο συγκεκριμένα, μία αύξηση – βελτίωση του δείκτη κατά 1.845 μονάδες φέρνει αύξηση μίας ποσοστιαίας μονάδας στις εισροές επενδύσεων. Όπως αναφέρθηκε και στην εισαγωγή της παρούσας εργασίας, οι χώρες την ανατολικής Ευρώπης πέρα από ομοιότητες έχουν και τεράστιες διαφορές μεταξύ τους, είτε αυτές είναι πολιτισμικές είτε κοινωνικό-πολιτικές. Σύμφωνα με σχετικούς δείκτες τέτοιες διαφορές είναι πολύ μεγαλύτερες στην συστάδα της ανατολικής Ευρώπης σε σχέση με την συστάδα του υπολοίπου. Σε αυτό το γεγονός έχουν παίξει πιθανότατα ρόλο και θεσμοί όπως η Ευρωπαϊκή Ένωση και η Ευρωζώνη, στην οποία συμμετέχει η συντριπτική πλειοψηφία της συστάδας του υπολοίπου της Ευρώπης σε αντίθεση με τις χώρες της άλλης συστάδας. Για παράδειγμα στην “ανατολική” συστάδα παρατηρούνται ζευγάρια όπως Ουκρανία – Εσθονία ή Αλβανία – Σλοβενία, οι διαφορές των οποίων στους περισσότερους τομείς είναι τεράστιες. Στην άλλη συστάδα υπάρχουν μεν διαφορές ανάμεσα στις χώρες αλλά όχι σε τέτοιο βαθμό, αφού οι περισσότερες λόγω και των θεσμών έχουν προσαρμοστεί σε μία σειρά όμοιων κανόνων. Αυτός ίσως να είναι και ο λόγος που η σημασία μεταβλητών που αντιπροσωπεύουν την πολιτική σταθερότητα και τα επίπεδα βίας στην κοινωνία να γίνεται πιο εμφανής σε εκείνη την συστάδα στην οποία οι διαφορές μεταξύ των χωρών είναι μεγαλύτερες. Άλλωστε ένας, από τους λόγους που επελέγη η ανατολική Ευρώπη για ανάλυση είναι αυτές οι εσωτερικές διαφορές των χωρών της συστάδας.

Χρησιμοποιούνται δηλαδή οι διαφορές ως ένα εργαλείο για γίνεται πιο εμφανής η σημασία της κάθε μεταβλητής.

Παρά την εισαγωγή της τρίτης μεταβλητής ελέγχου POLSTAB, τα αποτελέσματα-συμπεράσματα σχετικά με τις άλλες μεταβλητές δεν αλλάζουν. Η μεταβλητή VOICE και ο δείκτης διαφθοράς Cri παραμένουν στατιστικά σημαντικοί σε επίπεδο σημαντικότητας 1% για τις χώρες της ανατολικής Ευρώπης, όπως δηλαδή και στη στήλη 1.

Καταλήγοντας για το αυτό σύνολο μεταβλητών ελέγχου, πέρα από τα συμπεράσματα που εξάχθηκαν σχετικά με αυτές, το κυριότερο εύρημα είναι η διαφορά ανάμεσα στις δύο συστάδες χωρών όσον αφορά την κύρια ανεξάρτητη μεταβλητή Cri. Για τις χώρες της δυτικής Ευρώπης, ο δείκτης διαφθοράς έπαψε να είναι στατιστικά σημαντικός μόλις αφαιρέθηκαν οι επιδράσεις της ετεροσκεδαστικότητας. Αντίθετα, για τις χώρες της ανατολικής Ευρώπης παρέμεινε στατιστικά σημαντικός ο δείκτης, όχι μόνο μετά από την αφαίρεση της ετεροσκεδαστικότητας αλλά ακόμα και μετά την εισαγωγή των τριών μεταβλητών ελέγχου. Παρατηρούμε δηλαδή, ότι οι επιρροή του επιπέδου διαφθοράς κάθε ανατολικής χώρας είναι ισχυρή και φαίνεται να είναι ιδιαίτερα επίμονη στα αποτελέσματα των μοντέλων που δημιουργήθηκαν, δίνοντας μάλιστα πάντα θετικούς συντελεστές (coefficients).

Στο επόμενο στάδιο της εργασίας το αρχικό μοντέλο του πίνακα 2 θα εμπλουτιστεί με μεταβλητές ελέγχου που σχετίζονται με δείκτες χρηματοοικονομικών επιδόσεων και ανθρώπινου κεφαλαίου των χωρών. Τα αποτελέσματα θα δοθούν και πάλι με ξεχωριστούς πίνακες για τις δύο συστάδες.

**Πίνακας 4α: Εισαγωγή μεταβλητών ελέγχου σχετικών με δείκτες
χρηματοοικονομικών επιδόσεων και το ανθρώπινο κεφάλαιο για το υπόλοιπο της
Ευρώπης**

Rest of Europe

VARIABLES	(1) FDI	(2) FDI	(3) FDI
Cpi	-1.953 (1.486)	-2.009 (1.537)	-2.077 (1.558)
DEBTGDP	0.091* (0.046)	0.109 . (0.057)	0,184* (0,089)
HUMC	137.217 (223.37)	218.948 (322.91)	134.450 (275,98)
OPTRADE		-0.159 (0.221)	-0,226 (0,256)
GDPPC			0,005 (0,002)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.069	0.074	0.082
Adj R-squared	0.008	0.010	0.016
Number of Countries	19	19	19

Notes: Fixed Effects are employed. OPTRADE refers to openness to trade. GDPPC refers to GDP per capita. DEBTGDP refers to government to GDP. HUMC refers to human capital index.

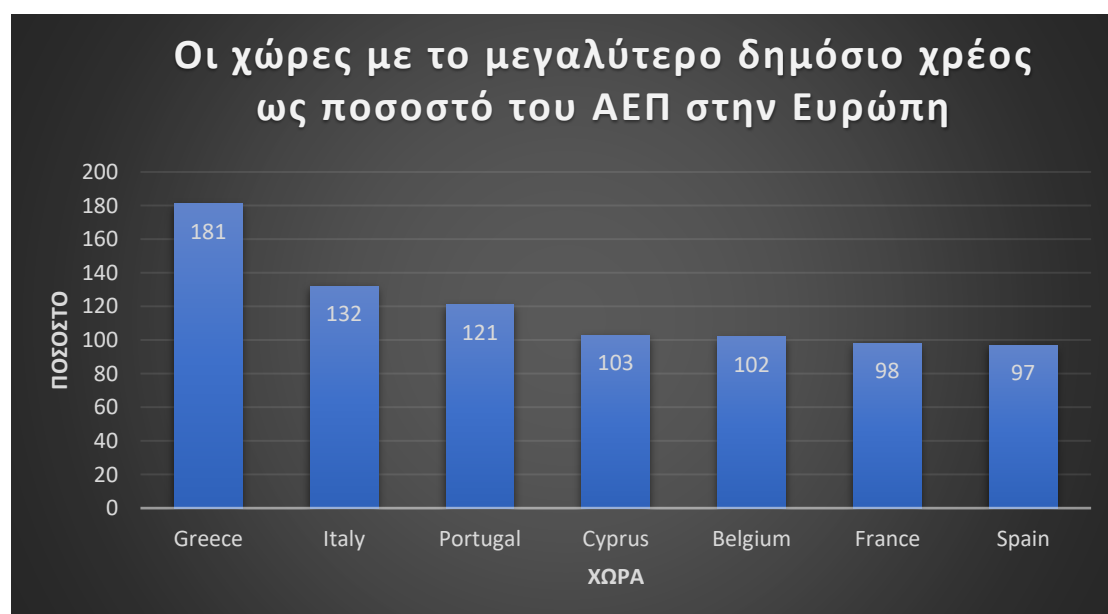
Robust standard errors in parentheses ***significant at 0 percent level, ** significant at 0.1 per cent level, * significant at 1 per cent level, . significant at 5 per cent

Στον πίνακα 4α δίνονται τα αποτελέσματα του μοντέλου μετά την εισαγωγή των νέων ανεξάρτητων μεταβλητών ελέγχου για τις χώρες του υπολοίπου Ευρώπης. Στην στήλη 1 εισάγονται δύο μεταβλητές ελέγχου μαζί με την κύρια ανεξάρτητη, η DEBTGDP και η HUMC. Η τελευταία αντιπροσωπεύει τη βαθμολόγηση του ανθρώπινου κεφαλαίου για κάθε χώρα σύμφωνα με το Διεθνές Νομισματικό Ταμείο. Η μεταβλητή αυτή δεν είναι στατιστικά σημαντική για τις άμεσες επενδύσεις σύμφωνα με τα αποτελέσματα του μοντέλου. Το πιο σημαντικό εύρημα της στήλης 1 και του πίνακα 4α είναι πως μετά την εισαγωγή του δείκτη DEBTGDP, ο οποίος αντιπροσωπεύει το δημόσιο χρέος κάθε χώρας ως ποσοστό του ΑΕΠ, η κύρια ανεξάρτητη μεταβλητή Cpi παραμένει όπως και στους προηγούμενους πίνακες που αφορούν το υπόλοιπο της Ευρώπης

στατιστικά μη σημαντική. Αντιθέτως, η μεταβλητή του χρέους για το μοντέλο της στήλης 1 είναι στατιστικά σημαντική για τις άμεσες ξένες επενδύσεις σε επίπεδο σημαντικότητας 1%.

Στη δεύτερη στήλη επαναλαμβάνεται το μοντέλο, μόνο που τώρα εισάγεται και η μεταβλητή OPTRADE, που συμβολίζει το πόσο ανοιχτή είναι μία χώρα στο εμπόριο και τις επενδύσεις. Για τις χώρες του υπολοίπου της Ευρώπης, η μεταβλητή αυτή φαίνεται να μην είναι στατιστικά σημαντική για την εξαρτημένη μεταβλητή FDI. Στη δεύτερη στήλη τα υπόλοιπα αποτελέσματα δεν αλλάζουν σε σχέση με την στήλη 1, καθώς η μεταβλητή που σχετίζεται με τη διαφθορά παραμένει μη σημαντική ενώ η μεταβλητή χρέους συνεχίζει να είναι στατιστικά σημαντική, εδώ όμως σε επίπεδο σημαντικότητας 5%. Όσον αφορά τη στήλη 3, το μοντέλο εμπλουτίζεται ακόμα περισσότερο με την εισαγωγή της μεταβλητής GDPPC, η οποία αντιπροσωπεύει το κατά κεφαλήν ΑΕΠ για κάθε χώρα. Για το μοντέλο που επικεντρώνεται στην συστάδα χωρών του υπολοίπου Ευρώπης, η μεταβλητή αυτή δεν είναι στατιστικά σημαντική για κανένα επίπεδο εμπιστοσύνης που θα μπορούσε να ληφθεί υπόψη. Ομοίως με τη στήλη 2, τα αποτελέσματα που αφορούν την κύρια ανεξάρτητη μεταβλητή και τις άλλες μεταβλητές ελέγχου παραμένουν ίδια, καθώς η Cpi δεν φαίνεται να παίζει κάποιο ρόλο στις αποφάσεις των επενδυτών για αυτό το σύνολο χωρών, ενώ η μεταβλητή χρέους συνεχίζει να είναι σημαντικός παράγοντας και μάλιστα για $\alpha = 1\%$.

Γράφημα 10: Χώρες με υψηλότερα δημόσια χρέη στην Ευρώπη (%ΑΕΠ)



Στον πίνακα 4β παρουσιάζονται τα αποτελέσματα του εμπλουτισμένου με τις νέες μεταβλητές ελέγχου μοντέλου για την συστάδα που επικεντρώνεται στις χώρες της ανατολικής Ευρώπης. Το σημαντικότερο εύρημα του πίνακα αυτού είναι ότι πλέον μετά την εισαγωγή της μεταβλητής χρέους στο μοντέλο (στήλη 1), η μεταβλητή Cpi γίνεται στατιστικά μη σημαντική, πρώτη φορά για την συγκεκριμένη συστάδα. Άξιο προσοχής αποτελεί το γεγονός πως παρότι δεν προκύπτουν στατιστικά σημαντικά αποτελέσματα για τη μεταβλητή Cpi, οι αντίθετοι συντελεστές (coefficients) στην σύγκριση των δύο συστάδων παραμένουν. Συγκεκριμένα, για τις χώρες της δυτικής Ευρώπης προκύπτουν θετικοί συντελεστές, ενώ για την ανατολική προκύπτουν αρνητικοί.

Τα αποτελέσματα των πινάκων 4α και 4β είναι αρκετά όμοια, πράγμα που σημαίνει πως οι επενδυτές των δύο συστάδων έχουν επίσης όμοια αντιμετώπιση ως προς τις μεταβλητές που αντιπροσωπεύουν το ανθρώπινο κεφάλαιο και τις χρηματοοικονομικές επιδόσεις. Πιο συγκεκριμένα, στη στήλη 1 και 2 παρατηρούμε πως η μεταβλητή DEBTGDP, που αφορά το δημόσιο χρέος της κάθε χώρας, είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%, κάτι που προέκυψε και για τις χώρες του υπολοίπου Ευρώπης. Αντίστοιχα, όμοια με την άλλη συστάδα είναι και τα αποτελέσματα που αφορούν την μεταβλητή HUMC που αφορά το ανθρώπινο κεφάλαιο, αφού και εδώ φαίνεται να μην παίζει ρόλο στην κατανομή των πόρων των επενδυτών.

Στην τρίτη στήλη, προκύπτει η μόνη διαφορά σε σχέση με τα προηγούμενα μοντέλα αλλά και σε σχέση με την άλλη συστάδα. Η διαφορά αυτή έχει να κάνει με την μεταβλητή OPTRADE, που αφορά το πόσο ανοιχτή είναι κάθε χώρα στις εμπορικές συναλλαγές και στις επενδύσεις. Παρατηρούμε πως στην στήλη αυτή, με βάση τη μεταβλητή αυτή απορρίπτουμε τη μηδενική υπόθεση σε επίπεδο σημαντικότητας 5% άρα καταλήγουμε πως αυτή η μεταβλητή είναι στατιστικά σημαντική για τα επίπεδα άμεσων ξένων επενδύσεων. Πιο συγκεκριμένα, για μία αύξηση 0.034 στην μεταβλητή αυτή, η εξαρτημένη μεταβλητή αυξάνεται κατά μία ποσοστιαία μονάδα.

**Πίνακας 4β: Εισαγωγή μεταβλητών ελέγχου σχετικών με δείκτες
χρηματοοικονομικών επιδόσεων και το ανθρώπινο κεφάλαιο για την ανατολική**

Ευρώπη

Eastern Europe

VARIABLES	(1) FDI	(2) FDI	(3) FDI
Cpi	-0.018 (0.120)	-0.006 (0.126)	-0,017 (0,123)
DEBTGDP	-0.060* (0.029)	-0.066* (0.028)	-0,070* (0,028)
HUMC	-38.096 (29.333)	-40.972 (31.055)	-30,568 (33,944)
OPTRADE		0.027 (0.021)	0,034 . (0,019)
GDPPC			-0,002 (0,001)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.053	0.055	0.058
Adj R-squared	0.009	0.010	0.011
Number of Countries	19	19	19

Notes: Fixed Effects are employed. OPTRADE refers to openness to trade. GDPPC refers to GDP per capita. DEBTGDP refers to government to GDP. HUMC refers to human capital index.

Robust standard errors in parentheses ***significant at 0 percent level, ** significant at 0.1 per cent level, * significant at 1 per cent level, . significant at 5 per cent

Όσον αφορά τα αποτελέσματα που προκύπτουν από τους άλλους δείκτες της στήλης, αυτά παραμένουν αμετάβλητα, αφού η μεταβλητή χρέους παραμένει στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%, ενώ η κύρια ανεξάρτητη μεταβλητή όπως και στις προηγούμενες στήλες είναι στατιστικά μη σημαντική, αλλάζοντας τα δεδομένα σε σχέση με τους προηγούμενους πίνακες, που αφορούσαν μεταβλητές ελέγχου σχετικές με τη πολιτική σταθερότητα και τις κοινωνικές ελευθερίες.

Στο σημείο αυτό, αφού έχουν αναλυθεί ξεχωριστά τα μοντέλα ανάλογα την κατηγορία που ανήκει η κάθε μεταβλητή ελέγχου, θα είχε ενδιαφέρον να δημιουργηθεί ένα πλήρες μοντέλο, το οποίο θα περιέχει την κύρια ανεξάρτητη μεταβλητή Cri αλλά και τις υπόλοιπες μεταβλητές ελέγχου που επελέγησαν σε αυτή την ανάλυση. Στον πίνακα 5α δίνεται το πλήρες μοντέλο για το υπόλοιπο της Ευρώπης. Όπως παρατηρούμε, η μεταβλητή Cri παραμένει στατιστικά μη σημαντική και άρα δεν μπορούμε να απορρίψουμε τη μηδενική υπόθεση για κανένα επίπεδο σημαντικότητας. Όσον αφορά τις μεταβλητές ελέγχου τα αποτελέσματα είναι ακόλουθα με αυτά που προέκυψαν στους προηγούμενους πίνακες. Πιο συγκεκριμένα, η μεταβλητή $DEBTGDP$ που συμβολίζει το δημόσιο χρέος ως ποσοστό του ΑΕΠ για κάθε χώρα είναι και πάλι στατιστικά σημαντική. Αυτή η ανεξάρτητη μεταβλητή είναι ίσως αυτή με τις ισχυρότερες επιδράσεις στις εισροές ξένων επενδύσεων, και μάλιστα αυτό συμβαίνει και στις δύο συστάδες. Για το συγκεκριμένο μοντέλο, για την μεταβλητή του χρέους απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 5% και *ceteris paribus* καταλήγουμε πως για μία ποσοστιαία αύξηση μίας μονάδας οι εισροές FDI αυξάνονται κατά 0,248 της μονάδας.

Πίνακας 5α: Πλήρες μοντέλο για το υπόλοιπο της Ευρώπης

Rest of Europe

<u>VARIABLES</u>	<u>FDI</u>
Cpi	-2,049 (1,443)
DEBTGDP	0,248 . (0,134)
HUMC	183,11 (304,1)
OPTRADE	-0,163 (0,186)
GDPPC	0,002 (0,001)
ROL	49,147 . (30,60)
VOICE	-4,361 (7,956)
POLSTAB	1,038 (4,647)
Fixed Effects	Yes
Observations	342
R-squared	0.121
Adj R-squared	0.049
<u>Number of Countries</u>	<u>19</u>

Notes: Fixed Effects are employed.

Robust standard errors in parentheses***significant
at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent,
. significant at 5 per cent

Όσον αφορά τις υπόλοιπες μεταβλητές ελέγχου που σχετίζονται με χρηματοοικονομικές επιδόσεις καμία εξ αυτών δεν φαίνεται να επηρεάζει τις εισροές επενδύσεων. Το ίδιο συμβαίνει και την μεταβλητή HUMC, που αφορά το ανθρώπινο κεφάλαιο κάθε χώρας.

Σε ότι έχει να κάνει με την ομάδα μεταβλητών που ασχολείται με την πολιτική σταθερότητα, τους νόμους και τις ελευθερίες, η μεταβλητή ROL συνεχίζει να είναι στατιστικά σημαντική για την εξαρτημένη μεταβλητή του μοντέλου σε επίπεδο 5%.

Για μία αύξηση μίας μονάδας στον δείκτη (min=-2,5/max=2,5) οι εισροές FDI αυξάνονται κατά 49,17 ποσοστιαίες μονάδες. Αυτό το αποτέλεσμα είναι πολύ σημαντικό καθώς δείχνει το πόσο σημαντικοί είναι οι νόμοι και η τήρηση τους για τους επενδυτές της δυτικής Ευρώπης. Οι υπόλοιπες σχετικές μεταβλητές δεν παρουσιάζουν κάποιο στατιστικά σημαντικό αποτέλεσμα.

Πίνακας 5β: Πλήρες μοντέλο για την ανατολική Ευρώπη

Eastern Europe

<u>VARIABLES</u>	<u>FDI</u>
Cpi	0,119 (0,092)
DEBTGDP	-0,075* (0,030)
HUMC	-25,693 (28,39)
OPTRADE	0,060* (0,023)
GDPPC	0,002 . (0,001)
ROL	4,889 . (2.789)
VOICE	7,270* (3.362)
POLSTAB	0,272 (1.031)
Fixed Effects	Yes
Observations	342
R-squared	0.121
Adj R-squared	0.049
<u>Number of Countries</u>	<u>19</u>

Notes: Fixed Effects are employed.

Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Στον πίνακα 5β καταγράφονται τα αποτελέσματα του πλήρους μοντέλου για την συστάδα της ανατολικής Ευρώπης. Το κύριο συμπέρασμα που αφορά την βασική υπόθεση αυτής της διπλωματικής εργασίας είναι πως ο δείκτης διαφθοράς Cpi, μετά την εισαγωγή των μεταβλητών ελέγχου, παραμένει στατιστικά μη σημαντικός, όπως και στον πίνακα που ενσωματώθηκαν οι μεταβλητές χρηματοοικονομικών επιδόσεων (4β). Καταλήγουμε λοιπόν, πως για κανένα επίπεδο σημαντικότητας α δεν μπορούμε να απορρίψουμε την μηδενική υπόθεση και άρα όσον αφορά αυτήν την συμβατική ανάλυση, η διαφθορά δεν φαίνεται να επηρεάζει τις εισροές άμεσων ξένων επενδύσεων ούτε σε αυτήν την συστάδα.

Όσον αφορά τις μεταβλητές ελέγχου που σχετίζονται με τις χρηματοοικονομικές επιδόσεις, η ανατολική συστάδα παρουσιάζει αρκετά διαφορετική εικόνα σε σχέση με την δυτική. Ο αριθμός αυτών που εμφανίζονται να είναι στατιστικά σημαντικές είναι μεγαλύτερος. Πιο συγκεκριμένα, η μεταβλητή χρέους DEBTGDP εμφανίζεται και πάλι στατιστικά σημαντική σε επίπεδο 1%, όπως και στην άλλη συστάδα. Για μία μονάδα μείωσης του δείκτη αυτού, οι εισροές άμεσων ξένων επενδύσεων αυξάνονται κατά 0,075 ποσοστιαίες μονάδες. Αυτό το νούμερο μπορεί να ακούγεται μικρό, αλλά δεν πρέπει να ξεχνάμε πως πρόκειται για ποσοστό που αντιστοιχεί σε δισεκατομμύρια ή εκατομμύρια ευρώ, ανάλογα και με το ύψος του συνόλου των εισροών. Παρόμοια αποτελέσματα παίρνουμε και για τη μεταβλητή OPTRADE, κάτι που δείχνει ότι οι επενδυτές προτιμούν χώρες που είναι ανοιχτές στο εμπόριο και στις επενδύσεις. Μάλιστα μία βελτίωση του δείκτη κατά μία μονάδα για $\alpha=1\%$, οι εισροές επενδύσεων αυξάνονται κατά 0,060 ποσοστιαίες μονάδες. Η τρίτη μεταβλητή που εμφανίζεται να έχει στατιστικά σημαντικό αποτέλεσμα είναι η GDPPC, που εκφράζει το κατά κεφαλήν ΑΕΠ. Πιο συγκεκριμένα, για αυτήν την ανεξάρτητη μεταβλητή απορρίπτουμε την μηδενική υπόθεση για $\alpha=5\%$ και καταλήγουμε πως είναι στατιστικά σημαντική για τις εισροές επενδύσεων.

Όσον αφορά το δεύτερο group μεταβλητών, δηλαδή αυτό που σχετίζεται περισσότερο με τις κοινωνικές συνθήκες, τους πολίτες και την νομοθεσία, προκύπτουν δύο μεταβλητές που αξίζουν σχολιασμού. Η μεταβλητή ROL που εκφράζει την εμπιστοσύνη των πολιτών στους νόμους αλλά και την εφαρμογή τους, είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 5%. Πιο συγκεκριμένα, για

αύξηση μίας μονάδας του δείκτη ROL, οι εισροές ξένων επενδύσεων αυξάνονται κατά 4,889 ποσοστιαίες μονάδες, αποτέλεσμα το οποίο εάν μεταφραστεί σε νομισματικές μονάδες γίνεται άμεσα κατανοητή η βαρύτητα της μεταβλητής αυτής. Επιπροσθέτως, η δεύτερη μεταβλητή του group, η οποία εμφανίζεται να επιδρά στις ροές επενδύσεων είναι η VOICE. Η μεταβλητή αυτή σχετίζεται όπως έχει αναφερθεί ήδη με τις κοινωνικές ελευθερίες, και την ελευθερία των μέσων ενημέρωσης. Είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας $\alpha=1\%$ και για αύξηση μίας μονάδος οι εισροές ξένων επενδύσεων αυξάνονται κατά 7,27 ποσοστιαίες μονάδες.

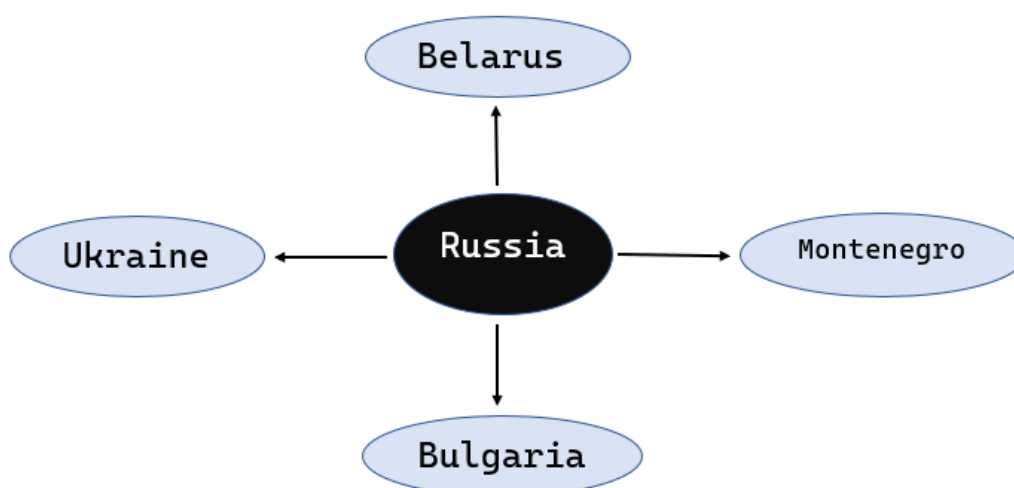
Οι διαφορές στα αποτελέσματα μεταξύ των δύο συστάδων για τα πλήρη μοντέλα είναι μεγάλες. Ο προφανής λόγος πίσω από αυτό είναι φυσικά οι διαφορές στα χαρακτηριστικά μεταξύ των χωρών των δύο συστάδων. Ένας μη προφανής λόγος όμως είναι και οι διαφορές των χωρών μέσα στην ίδια συστάδα. Δηλαδή υπάρχουν σημαντικές between αλλά και within διαφορές στα δεδομένα. Όπως έχει ήδη αναφερθεί, η συστάδα της ανατολικής Ευρώπης έχει μεγαλύτερο εύρος min/max στους δείκτες σε σχέση με την άλλη συστάδα. Για παράδειγμα μέσα στη συστάδα υπάρχει η Εσθονία, μία χώρα πρότυπο σε πολλούς τομείς, κυρίως στον χρηματοοικονομικό, αλλά υπάρχουν και χώρες όπως η Ουκρανία και η Αλβανία οι οποίες αντιμετωπίζουν σοβαρά προβλήματα σε πολλούς τομείς, είτε είναι θεσμικοί είτε οικονομικοί. Αυτός είναι και ο λόγος που επελέγη η ανατολική Ευρώπη σαν περιοχή έρευνας. Αυτές λοιπόν, οι within διαφορές ίσως να είναι εκείνες που κάνουν πιο ξεκάθαρα τα αποτελέσματα και φανερώνουν την σημασία ορισμένων δεικτών. Αντιθέτως, οι χώρες της συστάδας που αφορά κυρίως τη δυτική Ευρώπη, λόγω κοινών οργανισμών, θεσμών και κανόνων ίσως να λειτουργούν αντίστροφα. Δηλαδή όχι μόνο οι θεσμοί συγκρατούν συμπεριφορές όπως οι δωροδοκίες αλλά και όσον αφορά τα μοντέλα οι πολλές ομοιότητες ίσως να τα εμποδίζουν από το να αναδείξουν την σημασία των ανεξάρτητων μεταβλητών, από τη στιγμή που μιλάμε για panel δεδομένα.

7.2 Προσέγγιση home-host country

Σε αυτήν την ενότητα θα ακολουθηθεί μία διαφορετική προσέγγιση, η οποία, μετά τα αποτελέσματα των προηγούμενων μοντέλων, θα απομακρυνθεί για λίγο από τις

δύο συστάδες και θα επικεντρωθεί στις διμερείς σχέσεις μεταξύ των χωρών. Δεν είναι λίγα τα άρθρα που αναφέρουν πως πολλοί επενδυτές προτιμούν να επενδύουν σε χώρες, οι οποίες έχουν παρόμοια επίπεδα διαφθοράς με την χώρα καταγωγής τους, καθώς τέτοια περιβάλλοντα είναι οικεία και άρα οι στόχοι τους είναι πιο εύκολα επιτεύξιμοι. Πιο συγκεκριμένα, επελέγη η χώρα με το υψηλότερο ΑΕΠ και τις υψηλότερες εκροές FDI από κάθε συστάδα. Για την συστάδα Rest of Europe, η χώρα που επελέγη είναι η Γερμανία, ενώ για την συστάδα Eastern Europe η Ρωσία. Στην συνέχεια επελέγησαν τέσσερις χώρες για κάθε μία από αυτές τις δύο χώρες. Η επιλογή έγινε με βάση το που επιλέγουν οι Γερμανοί και οι Ρώσοι επενδυτές να αφιερώσουν τους πόρους τους. Η επιλογή έγινε κυρίως ως προς το μέγεθος του ποσοστού που καταλαμβάνει η Γερμανία και η Ρωσία ως επενδύτριες χώρες στο σύνολο εισροών των άλλων χωρών.

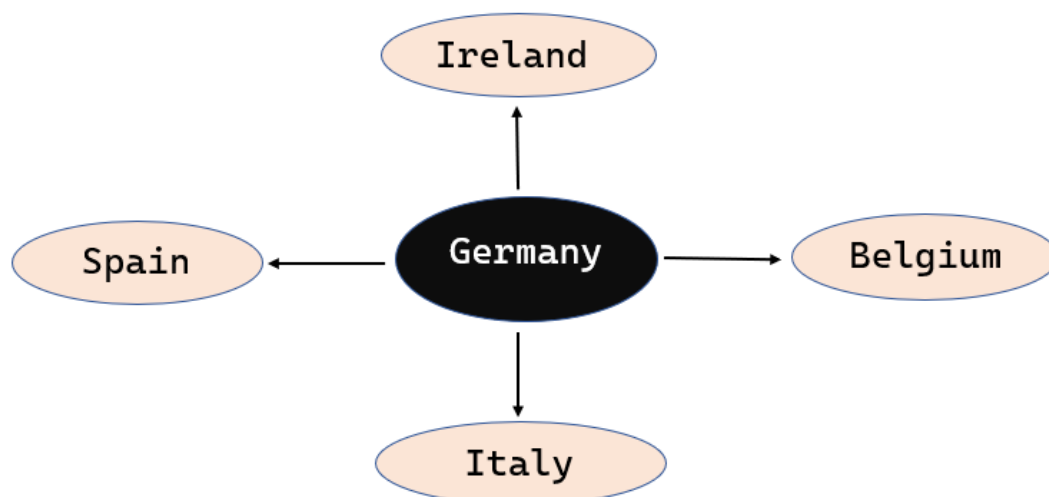
Γράφημα 11: Υποσυστάδα Ρωσίας.



Χώρες εκτός Ευρώπης, όπως για παράδειγμα η Κορέα, αγνοήθηκαν. Πιο συγκεκριμένα, home country στη πρώτη περίπτωση είναι η Ρωσία και host countries για εκείνη είναι η Ουκρανία, το Μαυροβούνιο, η Λευκορωσία και η Βουλγαρία. Αυτές είναι οι χώρες στις οποίες οι Ρώσοι επενδυτές καταλαμβάνουν το μεγαλύτερο μέρος ποσοστιαία ως προς το σύνολο των εισροών FDI. Στη δεύτερη περίπτωση, home country είναι φυσικά η Γερμανία και host countries η Ιρλανδία, η Ιταλία, το Βέλγιο

και η Ισπανία. Τα δεδομένα ως προς τα μεγέθη ροών κάθε χώρας συλλέχθηκαν από τον Ο.Ο.Σ.Α.

Γράφημα 12: Υποσυστάδα Γερμανίας.



Για το έτος 2020 η μέση διαφορά-απόσταση όσον αφορά τον δείκτη Cpi της home country από τις host είναι 14,75 για την τετράδα της Ρωσίας και 14,5 για την τετράδα της Γερμανίας. Κάτι που διαφοροποιεί τις δύο ομάδες είναι πως η Ρωσία έχει υψηλότερα επίπεδα διαφθοράς από τις τέσσερις χώρες που επενδύει, ενώ η Γερμανία έχει χαμηλότερα επίπεδα από τις δικές της επενδυτικές επιλογές.

Κάτι που γίνεται κατανοητό αν κοιτάξει κανείς τις δύο πεντάδες χωρών που δημιουργούνται με βάση τις εκροές της εκάστοτε home country είναι πως οι host countries βρίσκονται στην ίδια συστάδα με την home country. Αυτό ίσως να έχει να κάνει με όσα έχουν αναφερθεί στη βιβλιογραφία, δηλαδή το ότι οι επενδυτές ψάχνουν όμοια περιβάλλοντα. Όμως υπάρχει και ένας άλλος παράγοντας που πρέπει να αναφερθεί και αφορά κυρίως τη Ρωσία. Αυτός είναι πως τα τελευταία δύο χρόνια, έχουν αρχίσει και εντάσσονται όλο και περισσότερες ευρωπαϊκές χώρες σε συμφωνίες μείωσης των Ρωσικών επενδύσεων σε εδάφη της Ευρωπαϊκής Ένωσης, χώρες δηλαδή που βρίσκονται στην συστάδα Rest of Europe.

Στην προσέγγιση αυτή θα είχε ενδιαφέρον να φανεί πως επενδύουν οι “ηγέτες” των δύο συστάδων. Η Ρωσία είναι η χώρα όπου κατά μέσο όρο έχει τα υψηλότερα

επίπεδα διαφθοράς στα δεδομένα, ενώ η Γερμανία έχει αντίθετα από τα χαμηλότερα. Φυσικά κάποιος με μία επιφανειακή ανάλυση, θα μπορούσε να πει πως από τη στιγμή που οι δύο ηγέτιδες χώρες επιλέγουν χώρες με όμοια χαρακτηριστικά όσον αφορά τη διαφθορά, τα επίπεδα διαφθοράς είναι αυτά που αφορούν τους επενδυτές τους. Όμως αυτό είναι μία υπόθεση μόνο, όπως υπόθεση είναι και ότι τα επίπεδα FDI εξαρτώνται από την γεωγραφική απόσταση (*geographic proximity*), κάτι που επίσης έχει αναφερθεί στη βιβλιογραφία. Απάντηση δεν μπορεί να δοθεί αν δεν δημιουργηθούν μοντέλα πάνω σε συγκεκριμένο ερώτημα-υπόθεση. Το ερώτημα είναι ποιες μεταβλητές επηρεάζουν τους επενδυτές, οι οποίοι έχουν συνηθίσει σε ένα διεφθαρμένο περιβάλλον και ποιες εκείνους που έχουν συνηθίσει σε ένα περιβάλλον όπου τηρούνται απόλυτα οι νόμοι. Είναι η διαφθορά τελικά παράγοντας; Είναι τα αποτελέσματα διαφορετικά σε σχέση με τα πλήρη μοντέλα που αναπτύχθηκαν στην προηγούμενη ενότητα;

Τα αποτελέσματα για τα πλήρη μοντέλα που δημιουργήθηκαν για τις δύο ομάδες δίνονται στον παρακάτω πίνακα.

Πίνακας 6: Πλήρη μοντέλα για τις δύο ομάδες χωρών

VARIABLES	(1) GER/FD	(2) RUS/FD
Cpi	0,659 (0,592)	0,473** (0,140)
DEBTGDP	0,026 (0,159)	-0,06 (0,056)
HUMC	-169,62 (548,9)	-32,562 (62,093)
OPTRADE	0,282* (0,119)	0,094 (0,063)
GDPPC	-0,004 (0,008)	0,002*** (0,004)
ROL	23,437 (43,398)	27,61** (9,329)
VOICE	-0,989 (16,645)	16,159** (4,737)
POLSTAB	5,577 (5,771)	-0,171 (1,597)
Fixed Effects	Yes	Yes
Observations	71	71
R-squared	0.131	0.499
Adj R-squared	0.090	0.408
Number of Countries	4	4

Notes: Fixed Effects are employed. Robust standard errors in parentheses ***significant at 0 percent level, ** significant at 0.1 per cent level, * significant at 1 per cent level, . significant at 5 per cent

Στην πρώτη στήλη δίνονται τα αποτελέσματα για την ομάδα χωρών που επενδύει περισσότερο η Γερμανία. Τα αποτελέσματα, τα οποία προκύπτουν είναι αρκετά κοντά με αυτά που παρουσιάστηκαν στα μοντέλα των συστάδων. Το κυριότερο συμπέρασμα για αυτή την ομάδα είναι πως η μεταβλητή που αφορά την διαφθορά δεν εμφανίζεται να είναι στατιστικά σημαντική για κανένα α. Όσον αφορά τις

μεταβλητές ελέγχου, αυτή που φαίνεται να επηρεάζει τις ροές επενδύσεων είναι η μεταβλητή OPTRADE, η οποία συμβολίζει το πόσο ανοιχτή είναι κάθε χώρα στο εμπόριο και τις επενδύσεις. Για την μεταβλητή αυτή σε αυτό το μοντέλο απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 1% και σύμφωνα με τον συντελεστή της, για μία αύξηση μίας μονάδας οι εισροές FDI αυξάνονται κατά 0,282 ποσοστιαίες μονάδες.

Τα κυριότερα και πιο ενδιαφέροντα συμπεράσματα αυτού του πίνακα έρχονται στην δεύτερη στήλη, η οποία αφορά την ομάδα χωρών που αφιερώνουν τους πόρους τους οι Ρώσοι επενδυτές. Είναι η πρώτη φορά που μετά την εισαγωγή όλων των μεταβλητών ελέγχου, δηλαδή σε ένα πλήρες μοντέλο η μεταβλητή Cr1 της διαφθοράς είναι στατιστικά σημαντική και μάλιστα για $\alpha=0,1\%$. Σύμφωνα με το αποτέλεσμα αυτό με μία αύξηση μίας μονάδας του δείκτη, η εισροές FDI αυξάνονται κατά 0,473 ποσοστιαίες μονάδες. Πέρα όμως από την κύρια ανεξάρτητη μεταβλητή, παρατηρούμε πως αρκετές άλλες μεταβλητές είναι στατιστικά σημαντικές. Το κατά κεφαλήν ΑΕΠ (GDPPC) φαίνεται να επηρεάζει πολύ τις εισροές FDI, για αυτή την ομάδα χωρών επιβεβαιώνοντας την ανάγκη συμπερίληψης μεταβλητών σχετικών με τις χρηματοοικονομικές επιδόσεις στα μοντέλα που ερευνούν τη σχέση διαφθοράς και χρηματικών ροών.

Όσον αφορά το άλλο group μεταβλητών, όπως και στα μοντέλα που αφορούσαν τις χώρες της ανατολικής συστάδας στην προηγούμενη ενότητα, παρατηρούμε ότι προκύπτουν δύο στατιστικά σημαντικές μεταβλητές. Πιο συγκεκριμένα, για τη μεταβλητή ROL απορρίπτουμε τη μηδενική υπόθεση για $\alpha=0,1\%$ και σύμφωνα με το coefficient που προέκυψε, για αύξηση μίας μονάδας του δείκτη οι εισροές άμεσων ξένων επενδύσεων σε αυτές τις τέσσερις χώρες αυξάνονται κατά 27,61 ποσοστιαίες μονάδες. Παρόμοια αποτελέσματα προέκυψαν και για τη μεταβλητή VOICE, καθώς για μία αύξηση του δείκτη αυτού, η εξαρτημένη μεταβλητή FDI αυξάνεται κατά 16,159 ποσοστιαίες μονάδες (για $\alpha=0,1\%$).

Αυτό που επιβεβαιώνεται για ακόμα μία φορά μέσω του πίνακα είναι οι θετικοί συντελεστές που συναντάμε πάντα σε χώρες της ανατολικής συστάδας. Στην συγκεκριμένη περίπτωση, αυτός ο πίνακας δηλώνει πως οι Ρώσοι οι οποίοι επένδυσαν σε αυτές τις χώρες, για κάθε μοναδιαία αύξηση του δείκτη διαφθοράς

αυξάνουν κατά μισή περίπου ποσοστιαία μονάδα τις επενδύσεις τους. Οι θετικοί συντελεστές, δηλαδή η θετική σχέση μεταξύ διαφθοράς και FDI είναι κάτι που συναντάται αρκετά συχνά σε γεωγραφικές περιοχές με υψηλή διαφθορά και αυτό συνδέεται άμεσα με το γεγονός πως μία κατηγορία επενδυτών αναζητεί τέτοια περιβάλλοντα καθώς είναι γνώριμα με την home country τους, κάτι που αυξάνει τις πιθανότητες επίτευξης των στόχων τους. Σημαντικό επίσης είναι το αποτέλεσμα του συντελεστή προσαρμογής R-squared για την περίπτωση της ομάδας της Ρωσίας, ο οποίος αγγίζει το 50%, ένας αριθμός που για πραγματικά δεδομένα είναι αρκετά υψηλός.

Γίνεται λοιπόν κατανοητό πως η απάντηση στο αν η διαφθορά επηρεάζει τις ροές άμεσων ξένων επενδύσεων είναι δύσκολο να είναι μονολεκτική. Παράγοντες όπως η πηγή της επένδυσης, δηλαδή η χώρα από την οποία προέρχεται αυτή, αλλά και παράγοντες όπως ο χαρακτήρας του κάθε επενδυτή, η γεωγραφική τοποθεσία κ.α μπορούν να επηρεάσουν τα αποτελέσματα. Φαίνεται πως η ανάλυση των παραγόντων πρέπει να γίνεται κατά περίπτωση, καθώς είναι δύσκολο να βγει ένα και μόνο συμπέρασμα, το οποίο να αφορά όλα τα περιβάλλοντα, όλους τους επενδυτές και όλες τις γεωγραφικές περιοχές.

Η διπλωματική εργασία αυτή μέχρι στιγμής έχει επιβεβαιώσει την βιβλιογραφία για την περίπτωση των θετικών συντελεστών, των στατιστικά σημαντικών αποτελεσμάτων αλλά και την σημασία της πηγής της επένδυσης. Μπορεί να ακούγεται παράδοξο το γεγονός πως η διαφθορά σχετίζεται θετικά με τις ροές FDI, όμως είναι ένα αποτέλεσμα που συναντάται συχνά στην βιβλιογραφία και μάλιστα σε περιοχές με υψηλή διαφθορά, όπως δηλαδή η ανατολική Ευρώπη που πραγματεύεται η παρούσα εργασία. Η επιλογή της περιοχής αυτής και η σύγκριση της με την δυτική συστάδα δημιούργησε ενδιαφέροντα συμπεράσματα, τα οποία πηγάζουν από τις between και within διαφορές των συστάδων αυτών. Η ποικιλία των παραγόντων επίδρασης και οι διαφορές αυτές έχουν κρατήσει τους αναλυτές μακριά από αυτή την γεωγραφική περιοχή, καθώς ένα άρθρο-ανάλυση με μία σαφή απάντηση, μονολεκτικού χαρακτήρα, ίσως να τραβάει πιο πολύ την προσοχή. Όμως η ανάδειξη των παραγόντων αυτών, είτε είναι οικονομικοί είτε γεωγραφικοί είτε

θεσμικοί είναι πολύ σημαντικότερη στην προσπάθεια κατανόησης της συμπεριφοράς και των επιλογών των επενδυτών.

Η συντριπτική πλειοψηφία των εργασιών-άρθρων που στοχεύουν τη διαφθορά, είτε έχουν να κάνουν με χρηματικές ροές είτε όχι, χρησιμοποιούν σαν βάση τον δείκτη διαφθοράς Cpi. Ίσως αυτό το γεγονός να έχει εγκλωβίσει τους αναλυτές σε μία επαναλαμβανόμενη κατάσταση. Αυτό που ίσως θα μπορούσε να βοηθήσει είναι εναλλακτικοί ή επικουρικοί δείκτες που θα λειτουργούν ως proxies της διαφθοράς και θα μπορούν και εκείνοι να συνεισφέρουν στα μοντέλα και στην μελέτη αυτού του τόσο περίπλοκου ζητήματος. Στην επόμενη ενότητα προτείνεται και διεξάγεται μία εναλλακτική ανάλυση, με χρήση διαφορετικών proxies για τη διαφθορά, η οποία θα βασίζεται στον βιομηχανικό παράγοντα.

7.3 Εναλλακτική προσέγγιση

Στη παρούσα υπο-ενότητα των αποτελεσμάτων θα παρουσιαστούν τα αποτελέσματα της εναλλακτικής μεθόδου έρευνας της διαφθοράς για τις δύο συστάδες. Αυτή η μέθοδος θα βασιστεί σε δύο δείκτες της Παγκόσμιας Τράπεζας. Οι δείκτες αυτοί επελέγησαν καθώς περιέχουν πληροφορίες πάνω στους λεγόμενους διεφθαρμένους τομείς της οικονομίας, σύμφωνα με την έρευνα του οργανισμού Transparency International. Ο πρώτος δείκτης αφορά το πόσο συνεισφέρει η βιομηχανία στο ΑΕΠ κάθε χώρας. Ο δείκτης αυτός περιέχει τον εξορυκτικό, τον κατασκευαστικό, τον μεταποιητικό και τον ενεργειακό τομέα. Η μεταβλητή που αφορά τον συγκεκριμένο δείκτη ονομάστηκε IND. Ο δεύτερος δείκτης που θα χρησιμοποιηθεί αφορά την βιομηχανία αλλά χωρίς τον μεταποιητικό τομέα, καθώς ο τομέας αυτός επηρεάζει αρκετά τον πρώτο δείκτη και θα ήταν αρκετά ενδιαφέρον να ερευνηθεί μόνος του. Η μεταβλητή που περιέχει μόνο τον μεταποιητικό τομέα αφού αφαιρεθεί από τον πρώτο δείκτη ονομάζεται MAN. Το “υπόλοιπο” του δείκτη IND, δηλαδή μετά την αφαίρεση του MAN, θα ονομαστεί INDMM και φυσικά περιέχει τους υπόλοιπους τομείς που αναφέρθηκαν ως διεφθαρμένοι στην έρευνα. Όλες οι μεταβλητές δίνονται ως ποσοστό του ΑΕΠ

Table 7α: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή τον δείκτη Cpi για το υπόλοιπο της Ευρώπης

Rest of Europe

VARIABLES	Cpi	Cpi	Cpi
IND	-0.432 . (0.235)		
INDMM		-0.371 (0.244)	
MAN			-0.359 (0.316)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.070	0.022	0.036
Adj R-squared	0.014	0.035	0.021
Number of Countries	19	19	19

Notes: Fixed Effects are employed. IND refers to the industry value added including manufacturing sector as a percent of GDP. MAN refers to the manufacturing sector as a percent of GDP. INDMM refers to the Industry value added minus the manufacturing sector as a percent of GDP. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Πριν όμως ξεκινήσει η αντικατάσταση του δείκτη Cpi από τις μεταβλητές των διεφθαρμένων βιομηχανιών στο μοντέλο θα είχε ενδιαφέρον να φανεί πως αλληλοεπιδρούν εκείνες με τον συμβατικό δείκτη (Cpi). Στον πίνακα 5α, δημιουργούνται μοντέλα απλής γραμμικής παλινδρόμησης. Ως εξαρτημένη μεταβλητή τίθεται ο δείκτης διαφθοράς Cpi, ενώ σε κάθε στήλη του πίνακα χρησιμοποιείται διαφορετική ανεξάρτητη μεταβλητή, που φυσικά αντιστοιχεί κάθε φορά σε έναν ή περισσότερους διεφθαρμένους τομείς. Στην πρώτη στήλη η εξαρτημένη μεταβλητή IND αφορά το σύνολο των βιομηχανιών, δηλαδή συμπεριλαμβάνεται ο μεταποιητικός τομέας. Μάλιστα, για την συστάδα χωρών του υπολοίπου Ευρώπης η μεταβλητή φαίνεται να είναι στατιστικά σημαντική ως προς

τον δείκτη Cpi σε επίπεδο σημαντικότητας 5%. Όμως, εάν αυτός ο δείκτης περιέχει τις διεφθαρμένες βιομηχανίες, αυτό θα σήμαινε ότι η μεταβλητή θα έπρεπε να κινείται προς την ίδια κατεύθυνση με τον δείκτη Cpi. Πιο συγκεκριμένα θα περίμενε κανείς ότι μία αύξηση ως ποσοστό του ΑΕΠ αυτής της μεταβλητής θα έφερνε και αύξηση της διαφθοράς σε αυτή τη χώρα. Αντιθέτως, ο συντελεστής της IND είναι αρνητικός, πράγμα που σημαίνει πως η μείωση στη μεταβλητή αυτή φέρνει αύξηση στον δείκτη Cpi. Όσον αφορά τις στήλες 2 και 3 και τις αντίστοιχες μεταβλητές που χρησιμοποιήθηκαν επίσης δίνουν αρνητικούς συντελεστές (coefficients), όμως για καμία εκ των INDMM και MAN δεν προέκυψε στατιστικά σημαντικό αποτέλεσμα, ώστε να μπορεί να απορριφθεί η μηδενική υπόθεση.

Table 7β: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή τον δείκτη Cpi για την ανατολική Ευρώπη

Eastern Europe			
VARIABLES	Cpi	Cpi	Cpi
IND	1.334*** (0.357)		
INDMM		0.936* (0.406)	
MAN			1.135** (0.359)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.179	0.052	0.089
Adj R-squared	0.131	0.004	0.036
Number of Countries	19	19	19

Notes: Fixed Effects are employed. IND refers to the industry value added including manufacturing sector as a percent of GDP. MAN refers to the manufacturing sector as a percent of GDP. INDMM refers to the Industry value added minus the manufacturing sector as a percent of GDP. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Στον πίνακα 5β δίνονται τα αντίστοιχα αποτελέσματα για την συστάδα χωρών που αφορά την ανατολική Ευρώπη. Τα αποτελέσματα που προκύπτουν είναι αρκετά διαφορετικά σε σχέση με την άλλη συστάδα. Το πρώτο που παρατηρεί κανείς είναι

ότι οι συντελεστές και των τριών μεταβλητών είναι θετικοί, πράγμα που σημαίνει πως όσο αυξάνεται το ποσοστό που καταλαμβάνουν αυτοί οι διεφθαρμένοι τομείς στη “οικονομική πίτα” κάθε χώρας, τόσο χειροτερεύει και η βαθμολογία της χώρας σε ότι έχει να κάνει με τη διαφθορά. Αυτό το αποτέλεσμα ήταν και το αναμενόμενο εξαρχής, κάτι που όμως δεν επιβεβαιώθηκε για την συστάδα που αφορά το υπόλοιπο της Ευρώπης. Πιο συγκεκριμένα για την μεταβλητή IND, που αφορά το σύνολο των διεφθαρμένων βιομηχανιών, λαμβάνουμε ένα στατιστικά σημαντικό αποτέλεσμα σε επίπεδο σημαντικότητας πολύ κοντά στο 0. Σύμφωνα με τον συντελεστή, για μία αύξηση του δείκτη κατά μία μονάδα ποσοστιαία, ο δείκτης διαφθοράς αυξάνεται κατά 1.334 μονάδες.

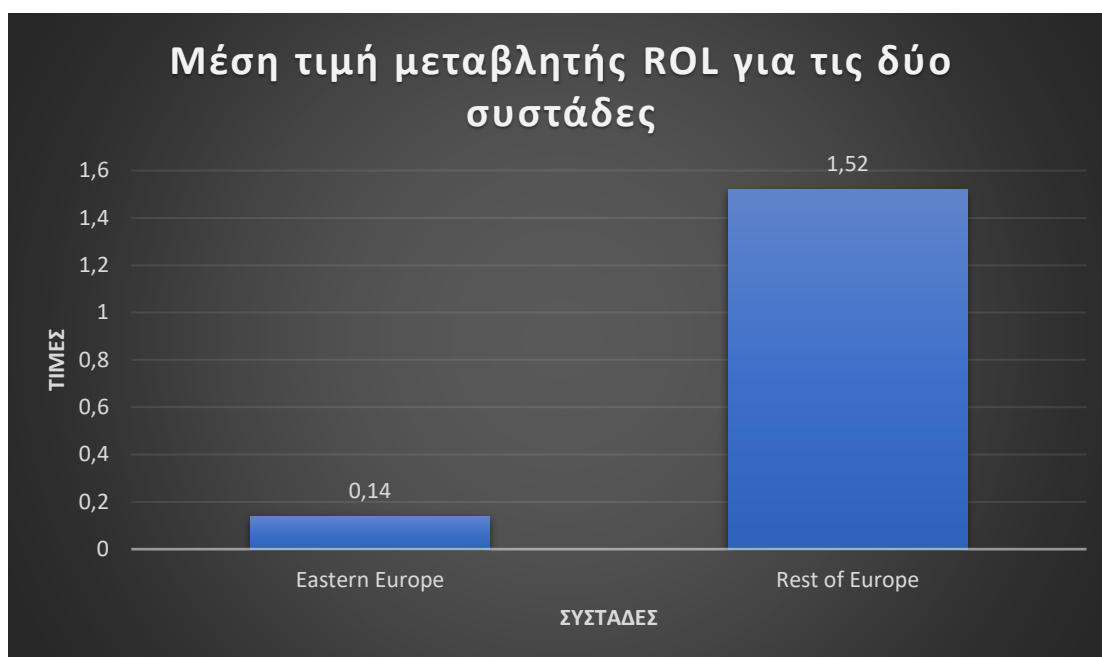
Ακόμα μία διαφορά που προκύπτει σε σχέση με την άλλη ομάδα χωρών είναι πως και τα παράγωγα της μεταβλητής IND, δηλαδή η MAN και η INDMM παρουσιάζουν στατιστικά σημαντικά αποτελέσματα. Η μεταβλητή που αντιπροσωπεύει το σύνολο των βιομηχανιών πλην του μεταποιητικό τομέα εμφανίζεται να είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%, και σύμφωνα με τον συντελεστή της μία αύξηση μία ποσοστιαία μονάδα οδηγεί σε αύξηση του δείκτη διαφθοράς κατά 0.936 της μονάδας. Αντίστοιχα, η μεταβλητή που αφορά τον μεταποιητικό τομέα μόνο, η MAN, είναι και αυτή στατιστικά σημαντική ως προς τον δείκτη διαφθοράς για $\alpha = 0.1\%$.

Αυτό που γίνεται για μία ακόμη φορά κατανοητό, μετά την ανάλυση των πινάκων 5α, 5β αλλά και των προηγούμενων, είναι πως τα αποτελέσματα αλλάζουν- εξαρτώνται σε μεγάλο βαθμό και από την γεωγραφική περιοχή που στοχεύει η εκάστοτε έρευνα. Στη βιβλιογραφία που σχετίζεται με την σχέση FDI – διαφθοράς γίνεται μεγάλη προσπάθεια στο να δοθεί μία και μοναδική απάντηση ως προς το αν υπάρχει σχέση μεταξύ τους και το αν αυτή είναι θετική ή αρνητική. Παρατηρούμε όμως πως οι συντελεστές των μεταβλητών διαφέρουν αρκετές φορές ανάλογα με την περιοχή στόχευσης. Αυτό σημαίνει πως ίσως η προσπάθεια να δοθεί μία σαφής απάντηση που να αφορά όλες τις περιοχές και όλες τις υπό-περιπτώσεις να μην είναι η σωστή προσέγγιση.

Το γεγονός πως παρατηρούμε διαφορετικά αποτελέσματα ανάμεσα στους πίνακες 5α και 5β είναι ιδιαίτερα ενδιαφέρον και μπορεί να σχετίζεται με αρκετούς

παράγοντες. Όπως γνωρίζουμε τα περισσότερα κράτη που ανήκουν στην συστάδα Rest of Europe ανήκουν σε θερμούς όπως Ε.Ε και Ευρωζώνη. Ο παράγοντας αυτός είναι πολύ σημαντικός, καθώς οι κανόνες-νόμοι που θέτουν τέτοιοι θεσμοί ακόμα και ως προϋπόθεση για την είσοδο σε αυτούς μπορεί να έχουν δημιουργήσει άμυνες στα συγκεκριμένα κράτη απέναντι στην ανεξέλεγκτη διαφθορά. Αυτό σημαίνει πως οι βιομηχανίες που αναφέρθηκαν στις συνεντεύξεις ως διεφθαρμένες, ίσως να αναγκάζονται να κινηθούν διαφορετικά σε ένα περιβάλλον που δημιουργεί εμπόδια σε τακτικές διαφθοράς. Πιο συγκεκριμένα, ακόμα και αν αυτές οι βιομηχανίες είναι όντως διεφθαρμένες, πρέπει να βρουν και τον ανάλογο χώρο-περιβάλλον για να αναδειχθούν αυτές οι συμπεριφορές. Αντίθετα, η πλειονότητα των χωρών της συστάδας που αφορά τις χώρες της ανατολικής Ευρώπης δεν ανήκουν σε αντίστοιχους διεθνείς οργανισμούς και ως εκ τούτου οι άμυνες τους απέναντι στη διαφθορά να είναι λιγότερες. Στον πίνακα 5γ παρουσιάζονται οι μέσοι όροι του δείκτη ROL για τις δύο συστάδες. Η μεταβλητή αυτή όπως ήδη αναφέρθηκε έχει ένα εύρος τιμών -2,5 με 2,5.

Γράφημα 13: Μέσος όρος μεταβλητής ROL για κάθε συστάδα



και αφορά το κατά πόσο κάθε κράτος είναι σε θέση να επιβάλλει τους νόμους στην κοινωνία, στην οικονομία αλλά και το πόση εμπιστοσύνη έχουν οι πολίτες του στο ίδιο το κράτος αλλά και στους νόμους. Όπως γίνεται κατανοητό από τον πίνακα 5γ, η ανατολική Ευρώπη έχει πολύ χαμηλότερο μέσο όρο από τις υπόλοιπες χώρες της Ευρώπης, συγκεκριμένα κατά 27,6% χαμηλότερο, κάτι που ίσως να επιβεβαιώνει την σημασία των θεσμών όπου συμμετέχουν οι χώρες της συστάδας Rest of Europe. Φαίνεται δηλαδή, ότι οι χώρες της ανατολικής Ευρώπης δυσκολεύονται περισσότερο να επιβάλλουν τους νόμους τους αλλά και ότι η εμπιστοσύνη των πολιτών στους θεσμούς των κρατών αυτών είναι λιγότερη. Τα παραπάνω δίνουν χώρο στη διαφθορά και στη περίπτωση μας στις διεφθαρμένες βιομηχανίες και ίσως να είναι αυτός ο λόγος για τον οποίο τα αποτελέσματα του πίνακα 5β να είναι τόσο διαφορετικά από αυτά του πίνακα 5α, ως προς τα στατιστικά σημαντικά αποτελέσματα.

Στον πίνακα 6α ξεκινάει η αντικατάσταση της μεταβλητής Cri, η οποία αποτέλεσε την κύρια ανεξάρτητη μεταβλητή των αρχικών μοντέλων. Η αντικατάσταση αυτή θα γίνει με τις μεταβλητές που αφορούν τις διεφθαρμένες βιομηχανίες με βάση την έρευνα και την συνέντευξη του οργανισμού Transparency International. Όπως και στα

αποτελέσματα της συμβατικής μεθόδου, έτσι και τώρα θα παρουσιαστούν πρώτα αυτά για το υπόλοιπο της Ευρώπης και στην συνέχεια αυτά που αφορούν την ανατολική Ευρώπη.

Με μία πρώτη ματιά στον πίνακα 6α, γίνεται εύκολα κατανοητό πως δεν προέκυψε κάποιο στατιστικά σημαντικό αποτέλεσμα για τις τρεις μεταβλητές. Είναι σημαντικό να αναφέρουμε πως αυτές οι μεταβλητές χρησιμοποιούνται σε ξεχωριστά μοντέλα, καθώς υπάρχουν υψηλές συσχετίσεις μεταξύ τους και αυτό θα μπορούσε να επηρεάσει τους κανόνες απόφασης για τους ελέγχους και να οδηγήσει σε λάθος συμπεράσματα. Κανένας εκ των δεικτών λοιπόν, δεν φαίνεται να επηρεάζει τις εισροές επενδύσεων, όσον αφορά τουλάχιστον τις χώρες του υπολοίπου της Ευρώπης.

Πίνακας 8α: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή την FDI για το υπόλοιπο της Ευρώπης

Rest of Europe			
VARIABLES	FDI	FDI	FDI
IND	-0.891 (1.438)		
INDMM		-1.469 (1.819)	
MAN			-0.333 (1.024)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.07	0.07	0.07
Adj R-squared	0.05	0.04	0.06
Number of Countries	19	19	19

Notes: Fixed Effects are employed. IND refers to the industry value added including manufacturing sector as a percent of GDP. MAN refers to the manufacturing sector as a percent of GDP. INDMM refers to the Industry value added minus the manufacturing sector as a percent of GDP. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

Για την συγκεκριμένη συστάδα αυτό το αποτέλεσμα είναι αρκετά αναμενόμενο. Όπως είδαμε και από τα μοντέλα που είχαν ως στόχο να δείξουν τις αλληλεπιδράσεις μεταξύ των νέων μεταβλητών και του δείκτη Cpi, δεν προέκυψε κάποια θετική ή αρνητική σχέση μεταξύ τους για τις χώρες αυτές. Αυτό με απλά λόγια σημαίνει ότι η αυξημένη ή μη παρουσία αυτών των βιομηχανιών δεν επηρεάζει την διαφθορά σε αυτές τις χώρες. Έχοντας δηλαδή ως δεδομένα τα αποτελέσματα του πίνακα 5α, δεν θα περιμέναμε κάποιο στατιστικά σημαντικό αποτέλεσμα στον πίνακα 6α. Ενισχύεται η άποψη με αυτό το αποτέλεσμα ότι οι χώρες του υπολοίπου της Ευρώπης είναι σε θέση μέσω των θεσμών και των κανόνων τους να χαλιναγωγήσουν την διαφθορά και τις επιδράσεις της στις εισροές ξένων επενδύσεων.

Στον πίνακα 6β που παρουσιάζεται παρακάτω απεικονίζονται τα αντίστοιχα αποτελέσματα για την ανατολική Ευρώπη. Συγκρίνοντας τον συγκεκριμένο πίνακα με τα αποτελέσματα του υπολοίπου Ευρώπης, παρατηρούμε μία τελείως διαφορετική εικόνα. Αντιθέτως, τα αποτελέσματα του 6β που αφορούν την εναλλακτική μέθοδο ανάλυσης της διαφθοράς επιβεβαιώνουν τα αποτελέσματα της συμβατικής μεθόδου ανάλυσης που προηγήθηκε, δηλαδή τους πίνακες 3β,4β. Πιο συγκεκριμένα, οι

Πίνακας 8β: Εισαγωγή μεταβλητών σχετικών με διεφθαρμένους τομείς έχοντας ως εξαρτημένη μεταβλητή την FDI για την ανατολική Ευρώπη

Eastern Europe			
VARIABLES	FDI	FDI	FDI
IND	0.566* (0.234)		
INDMM		0.284 (0.289)	
MAN			0.578* (0.232)
Fixed Effects	Yes	Yes	Yes
Observations	342	342	342
R-squared	0.03	0.06	0.02
Adj R-squared	0.03	0.05	0.04
Number of Countries	19	19	19

Notes: Fixed Effects are employed. IND refers to the industry value added including manufacturing sector as a percent of GDP. MAN refers to the manufacturing sector as a percent of GDP. INDMM refers to the Industry value added minus the manufacturing sector as a percent of GDP. Robust standard errors in parentheses***significant at 0 per cent level, **significant at 0.1 per cent level, *significant at 1 per cent, . significant at 5 per cent

συντελεστές που παρατηρούμε και στις τρεις στήλες-μεταβλητές είναι όλοι θετικοί. Για την μεταβλητή IND, που περιέχει όλες τις βιομηχανίες που θεωρήθηκαν διεφθαρμένες στην έρευνα, απορρίπτουμε την μηδενική υπόθεση σε επίπεδο σημαντικότητας 1% και καταλήγουμε πως αν το ποσοστό αυτών των βιομηχανιών στο ΑΕΠ των χωρών αυξηθεί κατά μία ποσοστιαία μονάδα, τότε οι εισροές ξένων επενδύσεων θα αυξηθούν κατά 0.566 ποσοστιαία μονάδα. Η μεταβλητή INDMM έχει μεν θετικό συντελεστή παραμένει όμως στατιστικά μη σημαντική και για αυτή τη συστάδα. Αντιθέτως, όσον αφορά την μεταβλητή που αντιπροσωπεύει τον μεταποιητικό τομέα, είναι στατιστικά σημαντική σε επίπεδο σημαντικότητας 1%, και συγκεκριμένα για αύξηση μίας ποσοστιαίας μονάδας στην ανεξάρτητη μεταβλητή αυτή, οι εισροές ξένων επενδύσεων αυξάνονται κατά 0.578 ποσοστιαίες μονάδες. Αυτό το εύρημα επιβεβαιώνει την έρευνα του οργανισμού Transparency International, καθώς σύμφωνα με τις συνεντεύξεις, αυτός ο τομέας κυριαρχεί σε συμπεριφορές όπως οι δωροδοκίες, ιδιαίτερα όταν πρόκειται για διεθνείς συμφωνίες, όπως είναι δηλαδή οι άμεσες ξένες επενδύσεις, αντικείμενο που πραγματεύεται φυσικά η παρούσα διπλωματική εργασία. Η μεταβλητή αυτή επελέγη να αφαιρεθεί από το σύνολο καθώς είναι ιδιαίτερα μεγάλο το ποσοστό που καταλαμβάνει συνήθως στα ΑΕΠ των χωρών. Η επιλογή αυτή φαίνεται πως ήταν επιτυχής, καθώς όπως παρατηρήθηκε η μεταβλητή INDMM, η μόνη δηλαδή που δεν περιέχει τον τομέα αυτόν δεν έδωσε στατιστικά σημαντικό αποτέλεσμα. Καταλήγουμε, πως κατά κύριο λόγο η μεταβλητή IND ήταν στατιστικά σημαντική ακριβώς επειδή περιείχε τον μεταποιητικό τομέα.

Τα αποτελέσματα της εναλλακτικής μεθόδου ανάλυσης της διαφθοράς και των εισροών επενδύσεων, επιβεβαιώνουν την συμβατική ανάλυση που προηγήθηκε. Για την συστάδα που αφορά το υπόλοιπο της Ευρώπης, οι επιδράσεις της διαφθοράς είναι ανύπαρκτες και στις δύο αναλύσεις. Σημαντικός φαίνεται να είναι ο ρόλος που παίζουν οι θεσμοί, οι νόμοι και η εμπιστοσύνη σε αυτούς. Για την συστάδα της ανατολικής Ευρώπης είναι πιο ισχυρές οι επιδράσεις της διαφθοράς μέχρι στιγμής και μάλιστα με θετικούς συντελεστές. Ένας από τους λόγους για αυτή τη θετική σχέση είναι πως υπάρχουν πολλοί επενδυτές που αναζητούν περιβάλλοντα με υψηλή διαφθορά και χαμηλά επίπεδα τήρησης των νόμων καθώς θεωρούν πως μέσω των

δωροδοκιών μπορούν να καταφέρουν ευκολότερα και γρηγορότερα στόχους που σε άλλα περιβάλλοντα δε θα μπορούσαν.

8. Συμπεράσματα

Στην παρούσα διπλωματική εργασία, μετά και την εκπόνηση του μοντέλου ARIMA και τις προβλέψεις που έδειξαν ισχυρή αύξηση των εισροών FDI για την δυτική και μηδαμινή για την ανατολική συστάδα, μελετήθηκε η σχέση διαφθοράς και εισροών άμεσων ξένων επενδύσεων με διάφορους τρόπους. Αρχικά, το βασικό μοντέλο που δημιουργήθηκε, δηλαδή πριν την εισαγωγή των μεταβλητών ελέγχου, έδειξε πως η μεταβλητή Cri που εκφράζει την διαφθορά είναι στατιστικά σημαντική για την ανατολική συστάδα και μόνο, κάτι που σημαίνει πως οι επενδύτες άμεσων ξένων επενδύσεων δίνουν σημασία στο επίπεδο διαφθοράς κάθε χώρας.

Μετά τον εμπλουτισμό του μοντέλου με τις μεταβλητές ελέγχου που σχετίζονται με την πολιτική σταθερότητα των χωρών και το επίπεδο κοινωνικών ελευθεριών που αυτές προσφέρουν στους πολίτες τους, τα συμπεράσματα δεν αλλάζουν καθώς ο συμβατικός δείκτης διαφθοράς Cri είναι στατιστικά σημαντικός μόνο για την συστάδα που αφορά τις ανατολικές χώρες της Ευρώπης. Όσον αφορά την συστάδα του υπολοίπου Ευρώπης η μεταβλητή παραμένει στατιστικά μη σημαντική. Ένα πολύ σημαντικό εύρημα είναι πως καθόλην την διάρκεια της εκπόνησης της εργασίας οι συντελεστές (coefficients) της μεταβλητής Cri των δύο συστάδων είναι αντίθετοι. Πιο συγκεκριμένα, για την συστάδα της δυτικής Ευρώπης οι συντελεστές είναι αρνητικοί, κάτι που φανερώνει μία αρνητική σχέση ανάμεσα στην διαφθορά και τις ροές επενδύσεων. Αντιθέτως, όσον αφορά την ανατολική συστάδα οι συντελεστές είναι θετικοί φανερώνοντας μια θετική σχέση ανάμεσα στην εξαρτημένη μεταβλητή και την κύρια ανεξάρτητη. Το τελευταίο είναι ένα εύρημα που συναντάται συχνά στην βιβλιογραφία που στοχεύει περιοχές με υψηλή διαφθορά όπως η Υποσαχάρια Αφρική. Όσον αφορά τις μεταβλητές ελέγχου που ενσωματώθηκαν στο μοντέλο, η μεταβλητή VOICE που αφορά τις ελευθερίες έκφρασης, είτε πρόκειται για τις

προσωπικές είτε για ελευθερία έκφρασης των Μέσων Μαζικής Ενημέρωσης είναι στατιστικά σημαντική και για τις δύο ομάδες χωρών. Η μεταβλητή που αντιπροσωπεύει την ποιότητα των θεσμών, ROL, είναι στατιστικά σημαντική μόνο για την συστάδα της δυτικής Ευρώπης, ενώ οι επενδυτές της ανατολικής Ευρώπης φαίνεται να ενδιαφέρονται περισσότερο για τις πολιτικές εξελίξεις και την σταθερότητα καθώς η μεταβλητή POLSTAB εμφανίζεται ως στατιστικά σημαντική όσον αφορά τις εισροές επενδύσεων. Γίνεται κατανοητό πως σύμφωνα με το μοντέλο αυτό για την ανατολική Ευρώπη η μεταβολή στις ροές επενδύσεων οφείλεται εν μέρει και στον δείκτη διαφθοράς.

Στο επόμενο στάδιο δημιουργήθηκαν μοντέλα όπου οι μεταβλητές ελέγχου αφορούσαν δείκτες χρηματοοικονομικών επιδόσεων και ανθρώπινου κεφαλαίου των χωρών. Αυτή η επιλογή αποδείχθηκε ιδιαίτερα σημαντική καθώς μετά την εισαγωγή των μεταβλητών ελέγχου είναι η πρώτη φορά όπου ο δείκτης διαφθοράς Crp εμφανίζεται στατιστικά μη σημαντικός για τις ροές επενδύσεων τόσο για την δυτική όσο και για την ανατολική συστάδα. Σύμφωνα με το μοντέλο αυτό λοιπόν γίνεται κατανοητό πως η μεταβολή στις ροές FDI οφείλεται στις μεταβλητές ελέγχου που επελέγησαν αυτή τη φορά και όχι στα επίπεδα διαφθοράς. Πιο συγκεκριμένα, και για τις δύο συστάδες φαίνεται να είναι ιδιαίτερα σημαντική για τους επενδυτές η μεταβλητή που αντιπροσωπεύει τον δημόσιο χρέος ως ποσοστό του ΑΕΠ, ενώ για την συστάδα των ανατολικών χωρών ιδιαίτερο ρόλο φαίνεται να παίζει και το πόσο ανοιχτή είναι κάθε χώρα σε επενδύσεις.

Στο τέλος της συμβατικής προσέγγισης δημιουργήθηκαν τα πλήρη μοντέλα και για τις δύο συστάδες. Όπως σε όλα τα μοντέλα της εργασίας χρησιμοποιήθηκαν Fixed Effects, αφαιρέθηκαν επιδράσεις της ετεροσκεδαστικότητας όπου αυτή ήταν παρούσα και φυσικά εκπονήθηκαν έλεγχοι για την ύπαρξη υψηλής συσχέτισης μεταξύ των μεταβλητών καθώς αυτό θα μπορούσε να επηρεάσει αποφάσεις σχετικές με την απόρριψη ή μη των υποθέσεων. Τα συμπεράσματα από τα πλήρη μοντέλα είναι ακόλουθα με τα προηγούμενα, καθώς ο δείκτης Crp της διαφθοράς παραμένει στατιστικά μη σημαντικός για τις δύο συστάδες, ενώ οι μεταβλητές ελέγχου που ήταν στατιστικά σημαντικές παραμένουν. Άξιο σχολιασμού είναι το γεγονός πως ο αριθμός μεταβλητών ελέγχου που εμφανίζονται να είναι στατιστικά σημαντικές είναι

μεγαλύτερος για την ανατολική συστάδα. Συγκεκριμένα, για την συστάδα του υπολοίπου Ευρώπης μόνο οι μεταβλητές που αφορούν το χρέος και την ποιότητα των θεσμών είναι στατιστικά σημαντικές, ενώ για την συστάδα των ανατολικών χωρών σχεδόν το σύνολο των χρηματοοικονομικών μεταβλητών, η ποιότητα των θεσμών και η μεταβλητή που αφορά τις ελευθερίες έκφρασης φαίνεται να επηρεάζουν τις εισροές άμεσων ξένων επενδύσεων.

Στην συνέχεια της ανάλυσης εκπονήθηκε η προσέγγιση home-host country. Σε αυτήν επελέγησαν από κάθε συστάδα οι χώρες με τις μεγαλύτερες εκροές επενδύσεων στον Ευρωπαϊκό χώρο. Συγκεκριμένα από τις δύο συστάδες επελέγησαν η Γερμανία και η Ρωσία. Αυτό έγινε για να αναλυθεί αν η προέλευση της επένδυσης επηρεάζει-διαφοροποιεί τα αποτελέσματα που βρέθηκαν στην συμβατική ανάλυση. Αυτό που έγινε κατανοητό είναι πως οι επενδυτές των χωρών αυτών φαίνεται να αφιερώνουν τους πόρους τους σε χώρες με όμοιο περιβάλλον με αυτό της χώρας από την οποία προέρχονται. Συγκεκριμένα, οι Γερμανοί επενδυτές φαίνεται να προτιμούν χώρες-μέλη της Ευρωζώνης με χαμηλή διαφθορά, ενώ οι Ρώσοι επενδυτές στοχεύουν περισσότερο ανατολικές χώρες με υψηλότερη διαφθορά και χαμηλότερη, σύμφωνα με τους δείκτες, ποιότητα θεσμών. Με βάση τις υποσυστάδες που δημιουργήθηκαν, τα νέα πλήρη μοντέλα έφεραν στο φως διαφορετικά αποτελέσματα σε σχέση με τα πλήρη μοντέλα των δύο αρχικών συστάδων. Σύμφωνα με τη νέα προσέγγιση, η σημαντικότερη διαφοροποίηση σχετίζεται με την υποσυστάδα της Ρωσίας. Συγκεκριμένα για αυτή την υποσυστάδα, η διαφθορά με βάση τον δείκτη Cpi είναι στατιστικά σημαντική για τις εισροές επενδύσεων παρά την εισαγωγή των μεταβλητών ελέγχου. Μάλιστα παρά το γεγονός πως παραμένουν σημαντικές οι μεταβλητές ROL, VOICE, GDPPC, ο δείκτης διαφθοράς είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 0.1%, ενώ άξιο αναφοράς είναι πως οι συντελεστές παραμένουν θετικοί. Αντιθέτως, για την υποσυστάδα της Γερμανίας τα αποτελέσματα είναι ακόλουθα με εκείνα της δυτικής συστάδας.

Στο τελευταίο μέρος της ανάλυσης εκπονήθηκε η εναλλακτική προσέγγιση που βασίστηκε πάνω στα στοιχεία του οργανισμού Transparency International σχετικά με τομείς της οικονομίας που παρατηρούνται συχνά δωροδοκίες και άλλες αντίστοιχες παράνομες ενέργειες και άρα μπορούν να χαρακτηριστούν ως διεφθαρμένοι. Ως

πρώτο εύρημα αυτού του μέρους της αναλυτικής διαδικασίας μπορεί να χαρακτηριστεί το γεγονός πως υπάρχει στατιστικά σημαντική σχέση ανάμεσα στην παρουσία αυτών των τομέων και στα επίπεδα διαφθοράς. Σύμφωνα με τα μοντέλα η αύξηση της παρουσίας των βιομηχανιών αυτών ως ποσοστό του ΑΕΠ μίας χώρας οδηγεί σε αύξηση του δείκτη διαφθοράς Cpi, και αυτό παρατηρείται πιο έντονα στην συστάδα της ανατολικής Ευρώπης. Στο δεύτερο σκέλος της εναλλακτικής προσέγγισης χρησιμοποιήθηκαν οι δείκτες των διεφθαρμένων βιομηχανιών ως εξαρτημένες μεταβλητές σε μοντέλα όπου η εξαρτημένη μεταβλητή ήταν οι εισροές FDI. Τα ευρήματα έχουν ιδιαίτερο ενδιαφέρον και φαίνεται να επιβεβαιώνουν την προσέγγιση home-host country. Συγκεκριμένα, για την ανατολική συστάδα και πάλι φαίνεται οι δείκτες των διεφθαρμένων τομέων να είναι στατιστικά σημαντικοί, με τον μεταποιητικό τομέα να επηρεάζει αρκετά τις εισροές άμεσων ξένων επενδύσεων. Αυτό το αποτέλεσμα είναι πολύ σημαντικό αφού κατά τη διάρκεια των συνεντεύξεων της έρευνας του οργανισμού Transparency International έγινε κατανοητό πως ο τομέας αυτός παρουσιάζει παρανομές ενέργειες και συμπεριφορές ιδιαίτερα όταν πρόκειται για ξένες συμφωνίες και επενδύσεις. Άρα τα ευρήματα της προσέγγισης επιβεβαιώνουν σε μεγάλο βαθμό αρκετά από τα στοιχεία της έρευνας του οργανισμού. Αξίζει να αναφερθεί πως και εδώ οι συντελεστές των νέων μεταβλητών είναι θετικοί, κάτι που δείχνει πως εάν μεγαλύτερο ποσοστό του ΑΕΠ καταλαμβάνεται από διεφθαρμένες βιομηχανίες, τότε οι εισροές ξένων επενδύσεων θα αυξηθούν. Το τελευταίο επίσης είναι ακόλουθο με τα ευρήματα της προσέγγισης home-host country, όπου έδειξε πως οι Ρώσοι επενδυτές προτιμούν χώρες με μεγάλη διαφθορά, δημιουργώντας και πάλι θετική σχέση ανάμεσα σε διαφθορά και εισροές άμεσων ξένων επενδύσεων. Όσον αφορά την δυτική συστάδα το ποσοστό που καταλαμβάνουν αυτοί οι τομείς στο ΑΕΠ κάθε χώρας δεν φαίνεται να επηρεάζουν τις εισροές FDI.

Η παρούσα διπλωματική εργασία προσπάθησε να δώσει απάντηση σε ένα ερώτημα που βασανίζει πολλούς αναλυτές, φορείς αλλά και κυβερνήσεις. Πως αντιδρούν οι επενδυτές απέναντι στη διαφθορά; Τα αποτελέσματα της εργασίας έδειξαν πως η χρήση του συμβατικού δείκτη Cpi που κυριαρχεί στην βιβλιογραφία δεν είναι αρκετή. Ίσως η δημιουργία proxies, όπως αυτά των βιομηχανιών που δημιουργήθηκαν κατά

τη διάρκεια της ανάλυσης να μπορούν να δώσουν μία άλλη εικόνα και επικουρικά να συνεισφέρουν για πιο ακριβή συμπεράσματα. Εκτός όμως από την ανάγκη εύρεσης και χρήσης εναλλακτικών εργαλείων, η εργασία ανέδειξε το ότι η απάντηση στο αν η διαφθορά επηρεάζει ή όχι, θετικά ή αρνητικά τις επενδύσεις, μπορεί να εξαρτάται από παράγοντες όπως τα κίνητρα του κάθε επενδύτη, την χώρα προέλευσης αλλά και την γεωγραφική περιοχή που εξετάζεται. Απάντηση μονολεκτική σε αυτό το ερώτημα λόγω της πολυπλοκότητας και των αριθμό των παραγόντων δεν μπορεί να υπάρξει, όμως αυτό δε σημαίνει ότι δεν χρειάζεται περαιτέρω ανάλυση. Αντιθέτως, είναι τέτοια η σημασία των άμεσων ξένων επενδύσεων, ιδιαίτερα για χώρες μικρού και μεσαίου εισοδήματος, που κάνει απαραίτητες αναλύσεις που θα στοχεύουν διαφορετικές γεωγραφικές περιοχές και τις εκάστοτε ιδιαιτερότητες τους. Περαιτέρω ανάλυση χρίζουν επίσης τομείς όπως τα κίνητρα βραχυπρόθεσμων και μακροπρόθεσμων επενδυτών, τι είδους δηλαδή επενδυτές έλκει κάθε χώρα-περιοχή αλλά επίσης και το πως μπορεί να επηρεάσει τα αποτελέσματα η ποσοστιαία συμμετοχή γηγενών επενδυτών.

9.Βιβλιογραφία

Abed, G. and Davoodi, H., 2001. *Corruption, structural Reforms and Economic Performance in the Transition Economies. IMF Working Paper. WP/00/132.*

Ahlquist, J., 2006. *Economic Policy, Institutions, And Capital Flows: Portfolio And Direct Investment Flows In Developing Countries. International Studies Quarterly (2006) 50, 681–704.*

Aidt, T., 2003. *Economic Analysis of Corruption: A Survey. The Economics Journal, Vol. 113, pp. 632-652.*

Akcay, S., 2006. *Corruption and human development. The Cato Journal 26(1):29-48.*

Alquist, R., Berman, N., Mukherjee, R. and Tesar, L., 2018. *Financial Constraints, Institutions, and Foreign Ownership. Journal of International Economics Volume 118, May 2019, Pages 63-83.*

Angelov, N., 2018. *Foreign Aid vs. Foreign Direct Investment: Which Leads to Equitable Benefits? 76th Annual MPSA Conference, International Political Economy Section, the Politics of Foreign Aid, Section 1.*

- Bailey, N., 2018. *Exploring the Relationship between Institutional Factors and FDI Attractiveness: A Meta-Analytic Review*. *International Business Review* Volume 27, Issue 1, February 2018, Pages 139-148
- Bardhan, P., 1997. *Corruption and Development: A Review of Issues*. *Journal of Economic Literature* 35: 1320-1346.
- Bayley, D.H., 1966. *The Effects of Corruption in a Developing Nation*, *Western Political Quarterly*. Vol. 19. N4.
- Belgibayeva, A. and Plekhanov, A., 2015. *Does corruption matter for sources of foreign direct investment?* *European Bank for Reconstruction and Development*. Working Paper No. 176.
- Bellos, S. and Subasat, T., 2012. *Governance and Foreign Direct Investment: a Panel Gravity Model Approach*. *International Review of Applied Economics*, Vol. 26, pp. 303-328.
- Blagojević, S. and Damijan, J., 2013. *The Impact of Corruption and Ownership on the Performance of Firms in Central and Eastern Europe*. *Ideas.repec.org*. Available at: <https://ideas.repec.org/a/taf/pocoec/v25y2013i2p133-158.html>
- Borensztein, E., De Gregorio, J. and Lee, J., 1998. *How Does Foreign Direct Investment Affect Economic Growth?* *Journal of International Economics* Volume 45, Issue 1, 1 June 1998, Pages 115-135.
- Brada, J., Drabek, Z., Mendez, J. and Perez, M., 2019. *National Levels of Corruption and Foreign Direct Investment*. *Journal of Comparative Economics* Volume 47, Issue 1, March 2019, Pages 31-49.
- Brooks, C. 2018. *Introductory Econometrics for Finance* by Chris Brooks, 2018, University of Cambridge ESOL.
- Costica, M., 2012. *Characteristics of FDI Process in Central and Eastern Europe*. *Journal: Procedia - Social and Behavioral Sciences* v62 (2012-10-24): 388-392
- Freckleton, M. and Wright, A., 2012. *Economic growth, foreign direct investment, and corruption in developed and developing countries*. *Journal of Economic Studies* vol. 39, issue 6, 639-652.
- Globerman, S. and D. Shapiro, A. 2002, "Global Foreign Direct Investment Flows: The Role of Governance Infrastructure", *World Development*, Vol. 30, pp. 1899-1919.
- Habib, M., Zurawicki, L., 2002 *Corruption and Foreign Direct Investment*. *J Int Bus Stud* 33, 291–307. <https://doi.org/10.1057/palgrave.jibs.8491017>
- Haggard, S., 1990 *Institutions and growth in East Asia*. *St Comp Int Dev* 38, 53–81 <https://doi.org/10.1007/BF02686328>.
- Hattari, R. and Rajan, 2011. *How Different Are FDI And FPI Flows? Does Distance Alter The Composition Of Capital Flows?* [Online] *Ideas.repec.org*. Available at: <https://ideas.repec.org/p/hkm/wpaper/092011.html>.
- Hermes, N. and Lensink, R., 2003. *Foreign Direct Investment, Financial Development and Economic Growth*. *Journal of Development Studies* 40(1):142-163.
- Lensink, R., Hermes, N. and Murinde, V., 2000. *Capital Flight and Political Risk*. *Journal of International Money and Finance* Volume 19, Issue 1, February 2000, Pages 73-92.

- Lensink, R. and Morrissey, O., 2006. *Foreign Direct Investment: Flows, Volatility, And The Impact On Growth**. *Review of International Economics* 14(3):478-493.
- Leff, N.H., 1964. *Economic Development through Bureaucratic Corruption American Behavioral Scientist* 8: 8-14. Reprint in A.J. Heidenheimer, M. Johnston and V.T. LeVine(Eds.), *Political corruption: A handbook*, 389-403, 1989. Oxford, Transaction Books.
- Leyes, C., 1965. *What is the Problem about Corruption? Journal of Modern African Studies*. Vol.3.N26
- Lui, F., 1985. *An equilibrium queuing model of bribery. Journal of Political Economy*, Vol. 93, pp. 760-781.
- Macrae, J., 1982. *Underdevelopment and the Economics of Corruption: A Game Theory Approach*, *World Development* 10(8): 677-87.
- Mauro, P., 1995. *Corruption and growth*, *The Quarterly Journal of Economics*, Vol.110 No.3, pp.681-712
- Rose-Ackerman, S., 1978. *Corruption: A Study in Political Economy*, N.Y.: Academic Press, 20.
- Sebastian-Andrei Labes, 2015. "Fdi Determinants in Brics," *CES Working Papers*, Centre for European Studies, Alexandru Ioan Cuza University, vol. 7(2), pages 296-308, August.
- Simon, D. and Eitzen, D., 1990. *Elite Deviance*, 3rd ed. Boston etc.: Allyn and Bacon.
- Stein, E. and Daude, C., 2002, *Institutions, Integration and the Location of Foreign Direct Investment*, *OECD Global Forum on International Investment. New Horizons and Policy Challenges for Foreign Direct Investment in the 21st Century: OECD Publishing*.
- Wei, S., 1997, "Why is Corruption So Much More Taxing than Tax? Arbitrariness Kills", *NBER Working Paper No.6255*.
- Wei, S., 2000, *How Taxing is Corruption on International Investors. The Review of Economics and Statistics*, Vol. 82, pp. 1-11.
- Wheeler, D. and Mody, A., 1992. *International Investment Location Decisions. Journal of International Economics*, vol. 33, issue 1-2, 57-76
- World Bank Database, [available at http://data.worldbank.org/](http://data.worldbank.org/)
- European Investment Bank, 2020. *Impact of FDI on economic growth*.
- European Investment Bank, 2017. *Wind of change: Investment in Central, Eastern and South Eastern Europe*.
- Eversheds Sutherland, 2016. *'Beneath the Surface' bribery and corruption report*.

Appendix

A. Χώρες ανατολικής συστάδας

1	Albania
2	Belarus
3	Bulgaria
4	Croatia
5	Czech Republic
6	Estonia
7	Hungary
8	Latvia
9	Lithuania
10	Moldova
11	Montenegro
12	North Macedonia
13	Poland
14	Romania
15	Serbia
16	Slovak Republic
17	Slovenia
18	Ukraine
19	Russia

B. Χώρες δυτικής συστάδας

1	Germany
2	Greece
3	Italy
4	France
5	Austria
6	Netherlands
7	Belgium
8	Denmark
9	Spain
10	Portugal
11	Switzerland
12	Luxembourg
13	United Kingdom
14	Ireland
15	Sweden
16	Norway
17	Finland
18	Iceland
19	Cyprus

