

**ΣΧΟΛΗ ΕΠΙΣΤΗΜΩΝ & ΤΕΧΝΟΛΟΓΙΑΣ ΤΗΣ ΠΛΗΡΟΦΟΡΙΑΣ  
ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ**

**Εφαρμογή μεθόδων ανάλυσης διακοπτόμενων  
χρονοσειρών (interrupted time series) σε δεδομένα  
covid**

**ΜΑΪΜΑΝΗ ΜΑΡΓΑΡΙΤΑ ΜΑΡΙΑ**

ΑΜ: 6190076

Επιβλέπουσα Καθηγήτρια  
Πεντελή Ξανθή

**Διπλωματική Εργασία**

που υποβλήθηκε στο Τμήμα Στατιστικής  
του Οικονομικού Πανεπιστημίου Αθηνών στο πλαίσιο του  
Προπτυχιακού Προγράμματος Σπουδών

Αθήνα

Σεπτέμβριος 2023



## ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω την καθηγήτρια μου κα. Ξανθή Πεντελή, Επίκουρη Καθηγήτρια του ΟΠΑ, για την σημαντική και άμεση καθοδήγηση της.



## ΠΕΡΙΛΗΨΗ

ΜΑΡΓΑΡΙΤΑ ΜΑΡΙΑ ΜΑΪΜΑΝΗ

**Εφαρμογή μεθόδων ανάλυσης διακοπτόμενων χρονοσειρών σε δεδομένα covid.**

Σεπτέμβριος 2023

Η παρούσα εργασία αναλύει και εφαρμόζει μεθόδους ανάλυσης χρονοσειρών για την κατανόηση της διασποράς της πανδημίας COVID-19 στην Ελλάδα, λαμβάνοντας υπόψη και τις παρεμβάσεις, λόγω των μέτρων αντιμετώπισης της πανδημίας, στις χρονοσειρές. Καθώς η πανδημία είχε περιόδους lockdown, απαγόρευσης μετακινήσεων και άλλες αλλαγές στη συμπεριφορά του πληθυσμού, οι παραδοσιακές μέθοδοι ανάλυσης χρονοσειρών μπορεί να μην είναι επαρκείς. Στην εργασία, εφαρμόζεται η μέθοδος της τμηματικής παλινδρόμησης για να ληφθούν υπόψιν οι παρεμβάσεις στη συμπεριφορά των δεδομένων COVID-19. Η εργασία περιλαμβάνει συλλογή και ανάλυση δεδομένων COVID-19 προσπαθώντας να αναδείξει τους παράγοντες που επηρέασαν στον περιορισμό της πανδημίας αλλά και να προβλέψει την εξέλιξη της χωρίς την εφαρμογή του δευτέρου lockdown.

## **ABSTRACT**

This thesis analysis and applies time series analysis methods to understand the dispersion of the COVID-19 pandemic, taking into account the interventions of the pandemic response measures on the time series. As the pandemic had periods of lockdowns, movement bans and other changes in population behaviour, traditional time series analysis methods may not be adequate. In this context, the segmented regression method is applied to account for intervention in the behavior of the COVID-19 data. The thesis includes collection and analysis of COVID-19 data trying to highlight the factors that influenced the containment of the pandemic and to predict its evolution without the application of the second lockdown.



# Περιεχόμενα

<b>1</b>	<b>Εισαγωγή</b>	<b>1</b>
<b>2</b>	<b>Υλικό και Μέθοδοι</b>	<b>3</b>
2.1	Δεδομένα και Επιπολασμός . . . . .	3
2.2	Μέθοδοι . . . . .	7
<b>3</b>	<b>Αποτελέσματα και Ανάλυση</b>	<b>13</b>
3.1	Περιγραφική Ανάλυση . . . . .	13
3.2	Στατιστική Ανάλυση μέσω Τμηματικής Παλινδρόμησης (Segmented Regression) . . . . .	18
3.2.1	Μοντέλο με εφαρμογή παρέμβασης . . . . .	20
3.2.2	Counterfactual Σενάριο . . . . .	27
<b>4</b>	<b>Συμπεράσματα</b>	<b>32</b>



# Κατάλογος Σχημάτων

2.1	Χρονοδιάγραμμα διαστήματος μελέτης . . . . .	4
2.2	Παρουσίαση της μεθόδου της τμηματικής παλινδρόμησης . . . . .	7
3.1	A:Εξέλιξη επιπολασμού κατά τη διάρκεια περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown. B: SD Error-Bars για τον επιπολασμό για κάθε χρονικό διάστημα . . . . .	15
3.2	Ποσοστό αλλαγής στην διαμονή στο σπίτι και στην κινητικότητα συγκριτικά με το σημείο αναφοράς. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown. . . . .	16
3.3	Εξέλιξη εμβολιαστικής κάλυψης κατα την διάρκεια της περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown. . . . .	16
3.4	Πλήθος Διενεργούμενων test κατά την διάρκεια της περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown. . . . .	17
3.5	Διάγραμμα συσχέτισης συντελεστή Pearson μεταξύ υποψήφιων ανεξάρτητων μεταβλητών. . . . .	18
3.6	(A)Σχέση μεταξύ επιπολασμού και χρόνου, (B)Σχέση μεταξύ επιπολασμού και ποσοστού εμβολιασμένων. . . . .	19
3.7	A:Κατάλοιπα έναντι προσαρμοσμένων Τιμών, B:QQ-Διάγραμμα των καταλοίπων, C: Διάγραμμα Αυτοσυσχέτισης των καταλοίπων, D:Διάγραμμα Μερικής Αυτοσυσχέτισης των καταλοίπων . . . . .	21

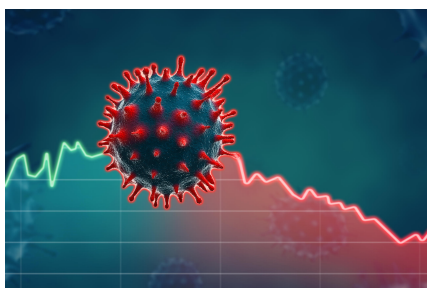
3.8	Διαγράμματα Αυτοσυσχέτισης και Μερικής αυτοσυσχέτισης τυποποιημένων καταλοίπων . . . . .	22
3.9	Τυποποιημένα Κατάλοιπα έναντι των ανεξάρτητων μεταβλητών . . . . .	23
3.10	Τυποποιημένα κατάλοιπα vs Χρόνου ανα επίπεδο Εφαρμογής Παρέμβασης . . . . .	23
3.11	A: Τυποποιημένα κατάλοιπα vs Προσαρμοσμένες τιμές B: QQ-plot τυποποιημένων καταλοίπων , C: Διάγραμμα αυτοσυσχέτισης, D: Διάγραμμα μερικής αυτοσυσχέτισης . . . . .	24
3.12	A: Κατάλοιπα vs Προσαρμοσμένες Τιμές, B: Διάγραμμα αυτοσυσχέτισης καταλοίπων, C: Διάγραμμα μερικής αυτοσυσχέτισης καταλοίπων . . . . .	27
3.13	A: Τυποποιημένα Κατάλοιπα έναντι προσαρμοσμένων τιμών του μοντέλου 3.4, B: QQ-plot τυποποιημένων καταλοίπων, C: Διάγραμμα Αυτοσυσχέτισης, D: Διάγραμμα Μερικής Αυτοσυσχέτισης . . . . .	28
3.14	Διάγραμμα διασποράς των μεταβλητών Επιπολασμός, Ημέρες απο το τέλος του πρώτου lockdown και αριθμό ημερίσιων test. . . . .	28
3.15	A: Τυποποιημένα Κατάλοιπα vs Προσαρμοσμένες Τιμές, B: QQ-plot τυποποιημένων καταλοίπων, C: Διάγραμμα Αυτοσυσχέτισης, D: Διάγραμμα Μερικής Αυτοσυσχέτισης . . . . .	29
3.16	Σύγκριση εξέλιξης επιπολασμού και counterfactual σενάριο . . . . .	30

# Κατάλογος Πινάκων

3.1	Περιγραφικά Μέτρα ποσοτικών μεταβλητών . . . . .	13
3.2	Υποψήφιος επεξηγηματικές μεταβλητές στο μοντέλο τμηματικής παλινδρόμησης με εξαρτημένη μεταβλητή τον επιπολασμό της νόσου . . . . .	19
3.3	Τιμές κριτηρίου AIC ανα τάξη $p$ και $q$ αυτοπαλίνδρομων κινητών μέσω μοντέλων . . . . .	22
3.4	Εκτιμώμενοι συντελεστές του μοντέλου τμηματικής παλινδρόμησης με εξαρτημένη μεταβλητή τον επιπολασμό της νόσου. . . . .	25
3.5	Αποτελέσματα μοντέλου (3.6) . . . . .	29
3.6	Αποτελέσματα τελικού μοντέλου(3.7) . . . . .	30
3.7	Συνοπτικός Πίνακας μεταβολής του επιπολασμού λόγω του lockdown . . . .	31

# Κεφάλαιο 1

## Εισαγωγή



Πηγή: Georgia Department of  
Public Health

<https://dph.georgia.gov/covid-19-status-report>

Οι κορονοϊοί είναι μια ομάδα στενά συγγενών ιών που προκαλούν ένα ευρύ φάσμα ασθενειών, από ένα κοινό κρυολόγημα έως πιο σοβαρές ασθένειες των τελευταίων ετών, όπως το SARS (Σοβαρό Οξύ Αναπνευστικό Σύνδρομο). Η νόσος του κοροναϊού (COVID-19) είναι μια μολυσματική ασθένεια που προκαλείται από τον ιό SARS-CoV-2. Είναι μία νέα μορφή κορονοϊού που δεν είχε καταγραφεί ποτέ μεταξύ άλλων γνωστών ιών, αναφέρθηκε για πρώτη φορά το 2019 και έκτοτε έχει εξαπλωθεί σε όλο τον κόσμο προκαλώντας 770.085.713 περιπτώσεις και νόσησης και 6.956.173 θανάτους. Η πανδημία έφτασε στην Ελλάδα τον Μάρτιο του 2020 και μέχρι σήμερα, 31/8/2023, έχουν μολυνθεί από τον ιό 5.352.647 άνθρωποι εκ των οποίων 5.315.336 κατάφεραν να αναρρώσουν. Οι περισσότεροι άνθρωποι που μολύνονται από τον ιό βιώνουν ήπια έως μέτρια αναπνευστική νόσο και αναρρώνουν χωρίς ειδική θεραπεία. Οι ηλικιωμένοι και εκείνοι με υποκείμενες ιατρικές παθήσεις όπως καρδιαγγειακή νόσο, διαβήτη, χρόνια αναπνευστική νόσο ή νεοπλασματικές ασθένειες έχουν περισσότερες πιθανότητες να αναπτύξουν σοβαρή ασθένεια.

Καθώς πολλές χώρες συνέχισαν να επηρεάζονται από το δεύτερο, τρίτο και τέταρτο κύμα του Covid-19, οι κυβερνήσεις με σκοπό να περιορίσουν την εξάπλωση της πανδημίας εφάρμοσαν μη-φαρμακευτικές παρεμβάσεις. Η Ελληνική Κυβέρνηση εφάρμοσε το πρώτο lockdown στις 9/3/2020 έως 4/5/2020. Η δεύτερη περίοδος lockdown οριοθετείται στο χρονικό διάστημα 5/11/2020 έως 14/5/2021, κατά τη διάρκεια της οποίας ξεκίνησε το πρώτο μαζικό πρόγραμμα εμβολιασμού (σχέδιο «Ελευθερία») το Δεκέμβριο του 2020.

Το lockdown αποτελεί μια μη φαρμακευτική παρέμβαση που εφαρμόζεται από την κυβέρνηση με σκοπό να περιορίσει την εξάπλωση μίας επιδημικής νόσου. Κατά την διάρκεια του lockdown οι μετακινήσεις εκτός σπιτιού είναι περιορισμένες με σκοπό να περιοριστούν οι δημόσιες επαφές και η εξάπλωση της νόσου. Αποτελεί μια πολιτική ελέγχου που περιλαμβάνει πολλές περιοριστικές πολιτικές, όπως το κλείσιμο των επιχειρήσεων, η υποχρεωτική χρήση μάσκας προσώπου, κοινωνική και φυσική απόσταση, απαγόρευση μαζικών εκδηλώσεων, διαδικτυακή και εξ αποστάσεως εκπαίδευση, περιορισμοί στα διεθνή και εσωτερικά ταξίδια.

Για την εφαρμογή και διατήρηση ενός lockdown πρέπει να λαμβάνονται υπόψιν τόσο τα πλεονεκτήματα όσο και τα μειονεκτήματα. Σύμφωνα με τους Woc-Colburn and Godinez (2022) όταν δεν υπάρχουν διαθέσιμα εμβόλια και θεραπείες, το lockdown συμβάλει σημαντικά στην μείωση της θνησιμότητας της νόσου καθώς και στην προστασία των ευπαθών ομάδων. Επίσης προσφέρει τον απαραίτητο χρόνο προκειμένου να αναπτυχθούν θεραπείες και φάρμακα. Ωστόσο ένα lockdown επιφέρει μια πληθώρα αρνητικών αποτελεσμάτων που είναι αντιληπτά τόσο στο παρόν όσο και στο μέλλον. Σύμφωνα με τους Lytras and Tsiodras (2021), με το lockdown είναι αρκετά δύσκολη η περίθαλψη χρόνιων και άλλων ασθενειών. Επίσης λόγω της απομόνωσης προκαλούνται ψυχικά προβλήματα στους ανθρώπους και υπάρχει μακροχρόνια αρνητική επίδραση στις μαθησιακές δεξιότητες των παιδιών λόγω του κλείσιμου των σχολείων. Φυσικά σημαντική είναι και η αρνητική επίδραση του lockdown στην οικονομία καθώς έτσι επηρεάζεται το δημόσιο σύστημα υγείας και επίσης αυξάνεται η ανεργία και η ανισότητα στην κοινωνία. Για τη διατήρηση ενός αυστηρού lockdown μία σημαντική αλλά αβέβαιη υπόθεση είναι η ανακάλυψη κάποιου φαρμάκου σε ποσότητα κατάλληλη να καλύψει τις ανάγκες του πληθυσμού αλλά και η εξασφάλιση ενός σημαντικού ποσοστού μη μολυσμένων ανθρώπων. (Lytras and Tsiodras, 2021)

Στην παρούσα εργασία, μέσω της στατιστικής μεθόδου των διακοπόμενων χρονοσειρών, στοχεύουμε να αξιολογήσουμε πόσο αποτελεσματικό ήταν το δεύτερο lockdown στον περιορισμό του Covid-19, καθώς και να εκτιμήσουμε την εξέλιξη των κρουσμάτων χωρίς την εφαρμογή του lockdown.

## Κεφάλαιο 2

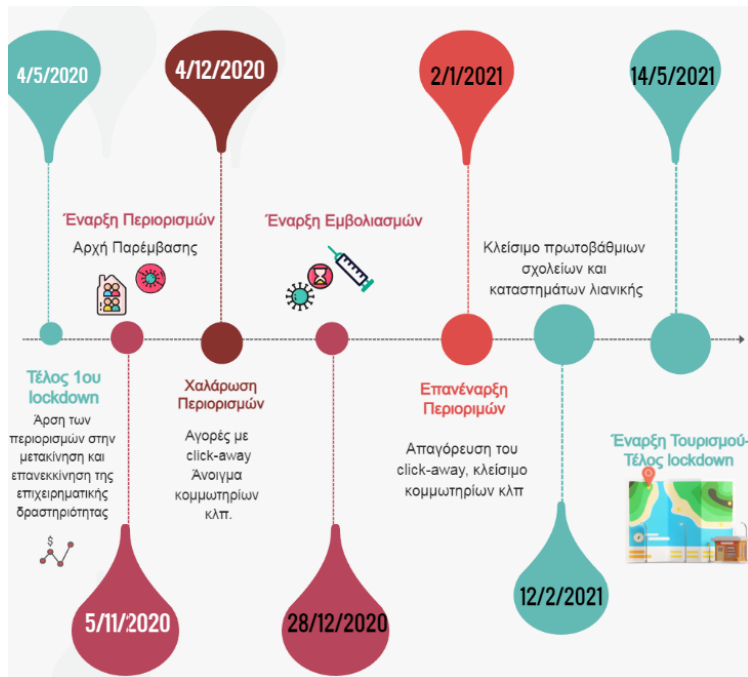
# Υλικό και Μέθοδοι

### 2.1 Δεδομένα και Επιπολασμός

Στην συγκεκριμένη εργασία χρησιμοποιήθηκε μία χρονοσειρά διάρκειας ενός χρόνου και 11 ημερών. Συγκεκριμένα, τα δεδομένα που θα αναλυθούν αναφέρονται στη χρονική περίοδο από το τέλος του πρώτου lockdown ως το τέλος του δεύτερου lockdown:

- Το πρώτο lockdown τερματίστηκε στις 4 Μαΐου 2020 όπου καταργήθηκε η γραπτή άδεια και τα SMS για τις μετακινήσεις .
- Το δεύτερο lockdown ξεκίνησε στις 5 Νοεμβρίου 2020 με την υποχρεωτικότητα του γραπτού μυνήματος για τις μετακινήσεις , την έναρξη της τηλεκαίτευσης και το κλείσιμο του λιανεμπορίου. Η άρση του lockdown ήρθε περίπου επτά μήνες μετά στις 15 Μαΐου 2021 με την κατάργηση του SMS και την έναρξη του τουρισμού.

Όπως φαίνεται στο Σχήμα 1, το δεύτερο lockdown σηματοδοτήθηκε από μερικές άρσεις περιορισμών και χαλάρωση των μέτρων, όπως επίσης και από την έναρξη του εμβολιασμού.



Σχήμα 2.1: Χρονοδιάγραμμα διαστήματος μελέτης

Τα δεδομένα συλλέχθηκαν από τον Εθνικό Οργανισμό Δημόσιας Υγείας (ΕΟΔΥ) και περιλαμβάνουν τον αριθμό των νέων test , κρουσμάτων και των εμβολιασμών με τουλάχιστον μία δόση καθώς επίσης και δεδομένα κινητικότητας. Στα δεδομένα που θα αναλυθούν περιελήφθη επίσης η παρουσία εμβολιαστικής κάλυψης και η εφαρμογή παρέμβασης (δίτιμες μεταβλητές). Επίσης υπολογίστηκε ο επιπολασμός όπως περιγράφεται. Πιο αναλυτικά :

► **Πλήθος νέων test στον πληθυσμό**

Αναφέρεται στο ημερήσιο πλήθος test που διενεργούνταν στον πληθυσμό, δεν γίνεται διαχωρισμός με βάση το είδος του test (rapid/μοριακό/self-test).

► **Πλήθος Νέων Κρουσμάτων**

Αποτελούν τα ημερησία επιβεβαιωμένα κρούσματα.

► **Ποσοστό Εμβολιασμών επί του Πληθυσμού**

Αποτελεί το ημερήσιο ποσοστό των εμβολιασμών με τουλάχιστον μία δόση στον πληθυσμό.

► **Επιπολασμός ανά 1.000.000 του πληθυσμού**

Ο επιπολασμός περιόδου είναι η αναλογία του πληθυσμού που έχει την εκάστοτε νόσο κατά τη διάρκεια μιας χρονικής περιόδου ενδιαφέροντος. Για τον υπολογισμό του επιπολασμού της νόσου Covid-19 επιλέξαμε το πλήθος νέων κρουσμάτων από τον πληθυσμό.

$$\text{Επιπολασμός Covid-19} = \frac{\text{Πλήθος κρουσμάτων}}{\text{Πληθυσμός}} * 1.000.000$$

Υπολογίστηκε βάσει της μεταβλητής "Πλήθος Νέων Κρουσμάτων". Άρα περιγράφει το πλήθος νέων περιπτώσεων της νόσου ανα ένα εκατομμύριο πληθυσμού.

Ο επιπολασμός είναι πολύ σημαντικός για τους σχεδιαστές συστημάτων υγείας και τους επαγγελματίες δημόσιας υγείας, γιατί αποτελεί ένα μέτρο της συχνότητας της νόσου στον πληθυσμό στο χρονικό διάστημα μελέτης. (What is prevalence?, n.d)

► **Ποσοστό αλλαγής της παραμονής στο σπίτι**

Την συγκεκριμένη μεταβλητή την λάβαμε από την αναφορά κινητικότητας της Google, (Covid-19 community mobility reports - google), είναι εκφρασμένη σε ποσοστό και εκφράζει τον τρόπο με τον οποίο άλλαξε η διάρκεια παραμονής στην κατοικία συγκριτικά με το σημείο αναφοράς. Το σημείο αναφοράς ορίζεται ως η μέση τιμή διάρκειας παραμονής στο σπίτι για την αντίστοιχη ημέρα της εβδομάδας κατά τη διάρκεια μιας περιόδου 5 εβδομάδων, από 3 Ιανουαρίου έως 6 Φεβρουαρίου 2020. Για παράδειγμα +26% σημαίνει ότι η παραμονή στο σπίτι είναι κατά 26% μεγαλύτερη σε σχέση με εκείνη της περιόδου από 3 Ιανουαρίου έως 6 Φεβρουαρίου 2020.

► **Ποσοστό αλλαγής της κινητικότητας για τόπους εκτός κατοικίας**

Αποτελεί το μέσο ποσοστό αλλαγής στις μετακινήσεις σε χώρους του λιανεμπορίου, φαρμακείων, συγκοινωνιών, χώρων εργασίας και πάρκων. Δηλαδή προστέθηκαν τα επιμέρους ποσοστά αλλαγής κινητικότητας τα οποία λήφθηκαν από τα δεδομένα κινητικότητας της Google και διαιρέθηκαν δια πέντε.

► **Παρουσία εμβολιαστικής κάλυψης**

Δίτιμη μεταβλητή που λαμβάνει την τιμή 1 από την ημερομηνία έναρξης του εμβολιασμού (2020-12-28, ημερομηνία έναρξης εμβολιασμού του υγειονομικού προσωπικού) έως το τέλος της περιόδου μελέτης, και 0 πριν την έναρξη.

► **Εφαρμογή παρέμβασης**

Η συγκεκριμένη μεταβλητή είναι και αυτή δίτιμη και λαμβάνει την τιμή 1 από την αρχή του δεύτερου lockdown έως τη λήξη του και 0 από τη λήξη του πρώτου lockdown μέχρι την έναρξη του δεύτερου.



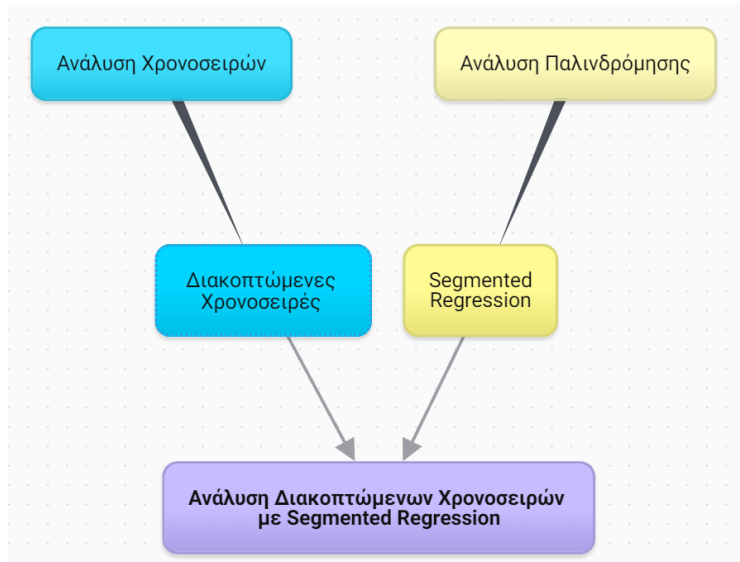
Πριν αναλυθούν οι μέθοδοι που θα εφαρμοστούν στην παρούσα εργασία, αξίζει να αναφερθεί ότι μέσα στο διάστημα μελέτης περιλαμβάνονται τρία κύματα του Covid-19 (Ατσαλάκης, 2021). Πιο αναλυτικά,

- 1ο κύμα: Στην χρονοσειρά περιλαμβάνεται δεδομένα ενός μικρού διαστήματος του πρώτου κύματος 4/5/2020-31/5/2020
- 2ο κύμα: Αποτελεί την πλειοψηφία των δεδομένων που θα αναλυθούν και περιλαμβάνει στις ημερομηνίες 1/6/2020-17/1/2021
- 3ο κύμα: από 18/1/2021 η Ελλάδα έρχεται αντιμέτωπη με το τρίτο κύμα του Covid-19 και έτσι τα υπόλοιπα δεδομένα αναφέρονται στο τρίτο κύμα, δηλαδή στις ημερομηνίες 18/1/2021- 15/5/2021.

Αξίζει να σημειωθεί ότι το κάθε κύμα ήταν πιο μεταδοτικό από το προηγούμενο. Πιο συγκεκριμένα τα ημερήσια κρούσματα ανέρχονταν στα 643 και 2.174 στο δεύτερο και τρίτο κύμα αντίστοιχα, ενώ κατά την διάρκεια του πρώτου κύματος τα ημερήσια κρούσματα ήταν 3 (Ατσαλάκης, 2021). Αυτό είναι σημαντικό να αναφερθεί καθώς οι εμβολιασμοί ξεκίνησαν στις 28/12/2020 όπου η Ελλάδα βρισκόταν στα μέσα του δευτέρου κύματος και συνεχίστηκαν κατά την διάρκεια του τρίτου.

## 2.2 Μέθοδοι

Στην παρούσα εργασία θα εφαρμοσθεί η μέθοδος της τμηματικής παλινδρόμησης (segmented regression) για την ανάλυση διακοπτόμενων χρονοσειρών (Interrupted Time Series). Στο παρακάτω σχήμα αναλύεται πως εφαρμόζεται η μέθοδος της τμηματικής παλινδρόμησης στην ανάλυση διακοπτόμενων χρονοσειρών.



Σχήμα 2.2: Παρουσίαση της μεθόδου της τμηματικής παλινδρόμησης

Η τμηματική παλινδρόμηση είναι μία μέθοδος που χρησιμοποιείται συχνά για την αξιολόγηση παρεμβάσεων όπως η εφαρμογή κυβερνητικών μέτρων σε συγκεκριμένο χρονικό διάστημα, η οποία είναι πιθανό να προκαλεί αλλαγή στην τάση και το επίπεδο μιας χρονοσειράς.

Η μέθοδος βασίζεται σε χρονοσειρές αποτελούμενες από παρατηρήσεις που λαμβάνονται από έναν πληθυσμό σε ίσα χρονικά διαστήματα. Στην παρούσα ανάλυση ως εξαρτημένη μεταβλητή χρησιμοποιείται η χρονοσειρά του ημερήσιου επιπολασμού, όπου διακόπτεται σε ένα συγκεκριμένο χρονικό διάστημα από την παρέμβαση. Προκειμένου λοιπόν να αξιολογηθεί η αποτελεσματικότητα της παρέμβασης αξιολογούνται δυο παράγοντες, αλλαγή στο επίπεδο και στην τάση της χρονοσειράς.

Αρχικό βήμα για την εφαρμογή της είναι ο διαχωρισμός της χρονοσειράς σε περιόδους πριν και μετά την παρέμβαση. Ως περίοδο πριν την παρέμβαση ορίζεται το χρονικό διάστημα πριν την εφαρμογή της παρέμβασης και αντίστοιχα ως μετά την παρέμβαση το χρονικό διάστημα μετά την εφαρμογή της.

Με την μέθοδο της τμηματικής παλινδρόμησης εφαρμόζεται η γραμμική παλινδρό-

μηση με την μέθοδο ελαχίστων τετραγώνων για κάθε χρονικό τμήμα. Πιο αναλυτικά η παλινδρόμηση έχει την παρακάτω βασική μορφή:

$$Y_t = \beta_0 + \beta_1 * t + \beta_2 * t \text{ after intervention} + \beta_3 * intervention_t + \varepsilon_t, \quad (2.1)$$

Όπου  $Y_t$  είναι η τιμή της εξαρτημένης μεταβλητής κατά την χρονική στιγμή  $t$  η οποία μπορεί να μετράται σε χρόνια, μήνες, μέρες κλπ, η μεταβλητή  $intervention_t$  λαμβάνει τις τιμές 1 και 0 αν κατά την χρονική στιγμή  $t$  υπάρχει εφαρμογή της παρέμβασης ή όχι αντίστοιχα. Η μεταβλητή  $t$  after intervention είναι ο χρόνος που έχει περάσει από την εφαρμογή της παρέμβασης και μετράται όπως η  $t$ . Τέλος τα  $\varepsilon_t$  είναι τα κατάλοιπα, δείχνουν την μεταβλητότητα που δεν έχει εξηγηθεί από το μοντέλο. Στο παραπάνω μοντέλο μπορούν να προστεθούν και άλλες επεξηγηματικές μεταβλητές.

Σε αυτό το μοντέλο, ο συντελεστής  $\beta_0$  εκφράζει το επίπεδο της μεταβλητής αποκρίσης κατα την αρχή της μελέτης. Ο συντελεστής  $\beta_1$  εκφράζει την αλλαγή στην μεταβλητή αποκρίσης πριν την έναρξη της παρέμβασης. Ο συντελεστής  $\beta_2$  εκφράζει την αλλαγή στην τάση της μεταβλητής ενδιαφέροντος μετά την έναρξη της παρέμβασης και τέλος ο  $\beta_3$  την αλλαγή στο επίπεδο αμέσως μετά την έναρξη της παρέμβασης. (Bernal et al.,2016)

Σχετικά με τα κατάλοιπα πρέπει να πληρούνται τρεις προϋποθέσεις:

1. Απουσία συσχέτισης μεταξύ των καταλοίπων,
2. Κανονικότητα καταλοίπων,
3. Ομοσκεδαστικότητα καταλοίπων.

Στην ανάλυση χρονοσειρών κατά την μοντελοποίηση της μεταβλητής  $Y_t$  είναι πιθανό να παραμείνει συσχέτιση μεταξύ των καταλοίπων, δηλαδή τα κατάλοιπα μίας χρονικής στιγμής  $t$  να επηρεάζονται από προηγούμενη χρονική στιγμή, όπως και ετεροσκεδαστικότητα, δηλαδή η διακύμανση των καταλοίπων ( $\sigma^2$ ) να μην είναι σταθερή στον χρόνο.

Για την ανίχνευση της αυτοσυσχέτισης των καταλοίπων χρησιμοποιείται ο έλεγχος Durbin-Watson καθώς και τα διαγράμματα αυτοσυσχέτισης (acf) και μερικής αυτοσυσχέτισης (pacf).

Προκειμένου να ελεγχθεί η ύπαρξη ετεροσκεδαστικότητας χρησιμοποιούνται διάφοροι έλεγχοι στην βιβλιογραφία. Στην παρούσα εργασία έγινε χρήση του ελέγχου White όπου θεωρεί ότι η διακύμανση των καταλοίπων είναι συνάρτηση των αρχικών ανεξάρτητων μεταβλητών η οποία περιλαμβάνει τις ανεξάρτητες μεταβλητές  $x_{ij}$ , τα τετράγωνα τους  $x_{ij}^2$  και τα μεταξύ τους γινόμενα  $x_{ij}x_{ik}$ ,  $i \neq k$ , δηλαδή

$$\hat{\varepsilon}_i^2 = h(\alpha_0 + \alpha_1 * x_{i1} + \dots + \alpha_p * x_{ip}) + v_i \quad (2.2)$$

Η παραπάνω παλινδρόμηση ονομάζεται βοηθητική παλινδρόμηση. Στην συνέχεια υλοποιείται ο παρακάτω έλεγχος:

$$H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_p = 0$$

$$H_1 : \exists j \in \{1, \dots, p\} : \alpha_j \neq 0$$

Η μηδενική υπόθεση αφορά την ομοσκεδαστικότητα των καταλοίπων και η εναλλακτική την ετεροσκεδαστικότητα. Το στατιστικό κριτήριο υπολογίζεται ως εξής:

$$LM_W = NR^2 \quad (2.3)$$

και ισχύει ότι  $LM_W \sim \chi^2_{(L)}$  όπου  $R^2$  είναι ο συντελεστής προσδιορισμού της βοηθητικής παλινδρόμησης και  $L$  οι βαθμοί ελευθερίας. Συγκεκριμένα αν η βοηθητική παλινδρόμηση περιλαμβάνει  $p+1$  μεταβλητές, μαζί με την σταθερά, ισχύει ότι

$$L = \frac{(p+1)((p+2))}{2} - 1 \quad (2.4)$$

(Τζαβαλής, 2008)

Όταν τα κατάλοιπα είναι ασυσχέτιστα και υπάρχει ομοσκεδαστικότητα, η διακύμανση των εκτιμητών  $\hat{\beta}$  υπολογίζεται ως εξής:

$$Var(\hat{\beta}) = \sigma^2(X'X)^{-1} \quad (2.5)$$

Όταν υπάρχει ετεροσκεδαστικότητα η διακύμανση των εκτιμητών δίνεται από την παρακάτω εξίσωση

$$Var(\hat{\beta}) = (XX')^{-1}X'\Omega X(X'X)^{-1} \quad (2.6)$$

όπου  $\Omega$  είναι ο πίνακας διακύμανσης-συνδιακύμανσης των καταλοίπων. Ο παραπάνω τύπος για την διακύμανση των εκτιμητών προέκυψε μέσω του ορισμού της διακύμανσης και της ιδιότητας της αμεροληψίας των εκτιμητών  $\beta$ . Πιο αναλυτικά οι εκτιμητές ελαχίστων τετραγώνων με την μορφή πινάκων γράφονται ως

$$\hat{\beta} = (X'X)^{-1}X'y \quad (2.7)$$

Αντικαθιστώντας στον παραπάνω τύπο την σχέση του πολλαπλού γραμμικού υποδείγμα-

τος  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$  προκύπτει ότι:

$$\begin{aligned}\hat{\boldsymbol{\beta}} &= (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}) = \\ &(\mathbf{X}'\mathbf{X})^{-1}(\mathbf{X}'\mathbf{X})\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon} = \\ &\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\end{aligned}\quad (2.8)$$

Παίρνοντας την αναμενόμενη τιμή της τελευταίας σχέσης προκύπτει ότι

$$\begin{aligned}E(\hat{\boldsymbol{\beta}}) &= E(\boldsymbol{\beta} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}) = \\ &E(\boldsymbol{\beta}) + E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}] = \boldsymbol{\beta}\end{aligned}\quad (2.9)$$

Έτσι αποδεικνύεται η ιδιότητα αμεροληψίας των εκτιμητών ελαχίστων τετραγώνων. Χρησιμοποιώντας τις παραπάνω σχέσεις στον ορισμό της διακύμανσης των εκτιμητών ελαχίστων τετραγώνων προκύπτει η σχέση 2.6, πιο αναλυτικά:

$$\begin{aligned}Var(\hat{\boldsymbol{\beta}}) &= E[(\hat{\boldsymbol{\beta}} - E(\hat{\boldsymbol{\beta}}))^2] = \\ &E[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta})'] = \\ &E[(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}'\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}] = \\ &(\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}')\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} = \\ &(\mathbf{X}\mathbf{X}')^{-1}\mathbf{X}'\boldsymbol{\Omega}\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\end{aligned}\quad (2.10)$$

(Τζαβαλής, 2008)

Στην παραπάνω σχέση  $E(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}') = \boldsymbol{\Omega}$  είναι ο πίνακας διακύμανσης-συνδιακύμανσης των καταλοίπων. Όταν τα κατάλοιπα παρουσιάζουν αυτοσυσχέτιση και ετεροσκεδαστικότητα ισχύει ότι  $\boldsymbol{\Omega} \neq \sigma^2\mathbf{I}$ , δηλαδή τα διαγώνια στοιχεία που αποτελούν την διακύμανση των καταλοίπων παύουν να είναι ίσα και τα μη διαγώνια που αποτελούν τις συνδιακυμάνσεις είναι διάφορα το μηδενός,  $\sigma_{ij} \neq 0$ . Έτσι ο εκτιμητής ελαχίστων τετραγώνων (LS) παραμένει αμερόληπτος αλλά δεν είναι πλέον αποτελεσματικός.

Προκειμένου να ληφθεί υπόψη η ύπαρξη ετεροσκεδαστικότητας και αυτοσυσχέτισης στην εκτίμηση των συντελεστών  $\boldsymbol{\beta}$  χρησιμοποιείται η Μέθοδος των Γενικευμένων Ελαχίστων Τετραγώνων (GLS) (Aitken, 1935)

$$\hat{\boldsymbol{\beta}}_{GLS} = (\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{X})^{-1}(\mathbf{X}'\boldsymbol{\Omega}^{-1}\mathbf{y})\quad (2.11)$$

Η μέθοδος των Γενικευμένων Ελαχίστων Τετραγώνων λαμβάνει υπόψιν αυτοπαλινδρομα μοντέλα (AR) και μοντέλα κινούμενου μέσου (MA), τα οποία χρησιμοποιούνται για την μοντελοποίηση της αυτοσυσχέτισης των καταλοίπων και η τάξη του κάθε μοντέλου αποφασίζεται από τα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης. Επίσης δίνεται η δυνατότητα να ληφθεί υπόψη και η ετεροσκεδαστικότητα. Για το σκοπό αυτό υπάρχουν διάφοροι μέθοδοι. Στην παρούσα εργασία χρησιμοποιείται η δομή της διακύμανσης που ονομάζεται "power of covariate" δομή διακύμανσης, όπου covariate είναι κάποια ανεξάρτητη μεταβλητή στο αρχικό μοντέλο. Πιο αναλυτικά, αφού κατασκευαστεί το διάγραμμα των τυποποιημένων καταλοίπων συναρτήσει των ανεξάρτητων μεταβλητών που χρησιμοποιούνται στην αρχική παλινδρόμηση και παρατηρηθεί ότι η διασπορά των καταλοίπων μεταβάλλεται όσο μεταβάλλονται και οι τιμές κάποιας ανεξάρτητης μεταβλητής  $X_i$ , η διακύμανση των καταλοίπων ( $\sigma^2$ ) πολλαπλασιάζεται με την απόλυτη τιμή της μεταβλητής  $X_i$  υψωμένη στην δύναμη  $2\delta$ , όπου η παράμετρος  $\delta$  εκτιμάται μέσω της profile πιθανοφάνειας (likelihood) (Andrzej Gałeccki, Tomasz Burzykowski, 2013, p.137). Έτσι τα κατάλοιπα ακολουθούν την παρακάτω κατανομή.

$$\varepsilon_i \sim N(0, \sigma^2 * |X_i|^{2\delta})$$

Στην περίπτωση που στην αρχική παλινδρόμηση υπάρχουν κατηγορικές μεταβλητές και παρατηρηθεί ότι το άπλωμα των καταλοίπων διαφέρει ανά επίπεδο της μεταβλητής η κατανομή των καταλοίπων αλλάζει σε

$$\varepsilon_{ij} \sim N(0, \sigma_j^2 * |X_{ij}|^{2\delta})$$

όπου j:επίπεδα της κατηγορικής μεταβλητής.

Η αλλαγή στην τάση (trend) και το επίπεδο (level) της χρονοσειράς μετά την εφαρμογή της παρέμβασης, αφού αντικατασταθούν στην εξίσωση 2.1 οι συντελεστές LS με τους GLS, θα αξιολογηθεί μέσω των συντελεστών  $\beta_2$  και  $\beta_3$ , αντίστοιχα. Η άμεση αλλαγή που υπέσχει η μεταβλητή ενδιαφέροντος μετά την παρέμβαση αξιολογείται από τον συντελεστή  $\beta_1$ . Αναλυτικότερα θα πρέπει να ερευνηθεί η σημαντικότητα των συντελεστών μέσω του παρακάτω ελέγχου

$$H_0 : \beta_i = 0 \text{ vs } H_1 : \beta_i \neq 0$$

Αναλόγως του προσήμου του συντελεστή θα κριθεί και αν η παρέμβαση ήταν αποτελεσματική. Πιο αναλυτικά αν η παρέμβαση είχε σκοπό να προκαλέσει μείωση στο επίπεδο και την τάση της μεταβλητής απόκρισης, για να κριθεί ως αποτελεσματική οι συντελεστές θα πρέπει να έχουν αρνητικό πρόσημο. Αντίστοιχα αν σκοπός ήταν η αύξηση στην τάση και το επίπεδο της μεταβλητής απόκρισης οι συντελεστές θα έπρεπε να έχουν θετικό πρόσημο. Στην παρούσα εργασία η παρέμβαση θα κριθεί ως αποτελεσματική αν οι παραπάνω συντελεστές έχουν αρνητικό πρόσημο.

## Κεφάλαιο 3

# Αποτελέσματα και Ανάλυση

### 3.1 Περιγραφική Ανάλυση

Το σύνολο δεδομένων που θα αναλυθεί στην παρούσα εργασία αποτελείται από 377 παρατηρήσεις και 8 μεταβλητές οι οποίες προαναφέρθηκαν στο Κεφάλαιο 2. Το πλήθος των παρατηρήσεων πριν και μετά την παρέμβαση είναι 185 και 192 αντίστοιχα. Οι μεταβλητές που περιέχονται στα δεδομένα διακρίνονται σε δυο κατηγορίες, ποσοτικές και ποιοτικές. Οι ποσοτικές μεταβλητές διακρίνονται σε συνεχείς και διακριτές. Στην πρώτη κατηγορία ανήκουν ο επιπολασμός, το ημερήσιο ποσοστό εμβολιασμένων, το ποσοστό εμβολιασμών επί του πληθυσμού, το ποσοστό αλλαγής της διαμονής στο σπίτι και αλλαγής κινητικότητας για τόπους εκτός κατοικίας, αντίστοιχα στην δεύτερη κατηγορία ανήκουν το πλήθος νέων κρουσμάτων και νέων test. Οι υπόλοιπες μεταβλητές είναι ποιοτικές και αποτελούν κατηγορικές μεταβλητές με δυο επίπεδα.

Στην συνέχεια παρουσιάζονται περιγραφικά μέτρα για τις ποσοτικές μεταβλητές για όλο το χρονικό διάστημα μελέτης.

	Ποσοστό αλλαγής της κινητικότητας για τόπους μη-κατοικίας	Ποσοστό αλλαγής της διαμονής στο σπίτι	Επιπολασμός ανά 1.000.00 Πληθυσμού	Ποσοστό εμβολιασμένων	Πλήθοςεν νέων test
Ελάχιστη Τιμή	-62.8	-8.0	0.19	0.0	106
1ο Ποσοστημόριο.	-24	1.0	12.0	0.0	6223
Διάμεσος	-2.6	6.0	53.9	0.0	11864
Μέση Τιμή	-2.3	6.0	95.4	4.7	12825
3ο Ποσοστημόριο.	14.6	11.0	159.0	4.4	18638
Μέγιστη Τιμή	67.6	25.0	414.0	40.9	63411
Τυπική Απόκλιση	26.0	6.4	100.6	9.2	0.1

Πίνακας 3.1: Περιγραφικά Μέτρα ποσοτικών μεταβλητών

Το αρνητικό πρόσημο στις δυο πρώτες μεταβλητές του Πίνακα 3.1 εκφράζει ότι οι άνθρωποι περνούν λιγότερο χρόνο στα συγκεκριμένα μέρη σε σχέση με το σημείο αναφοράς.



Άρα απο τον Πίνακα 3.1 παρατηρείται ότι η μέγιστη μείωση της κινητικότητας σε τόπους λιανεμπορίου ,φαρμακείων, συγκοινωνιών, χώρους εργασίας και πάρκων είναι κατα μέσο όρο 62.8% και αντίστοιχα η μέγιστη αύξηση είναι 67.6% ,επίσης στο 25% των ημερών η κινητικότητα σε αυτά τα μέρη παρουσιάζει κατα μέσο όρο μείωση μεγαλύτερη απο 24% και το 50% των ημερών έχουν μεγαλύτερη μείωση απο 2.6% .

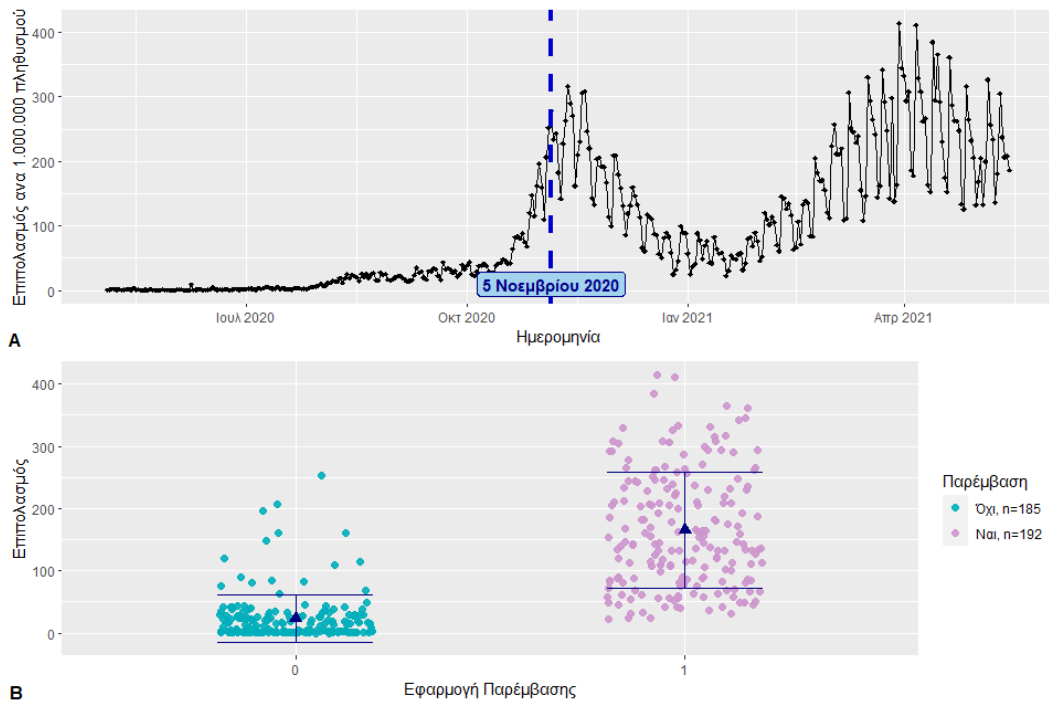
Αντίστοιχη ερμηνεία δίνεται και για το ποσοστό αλλαγής της διαμονής στο σπίτι όπου η μέγιστη μείωση που παρατηρήθηκε είναι 8% και η μέγιστη αύξηση 25% σε σχέση με το σημείο αναφοράς.

Είναι σημαντικό να γίνει μια διερεύνηση των ημερών όπου παρατηρήθηκαν οι πιο ακραίες τιμές στην κινητικότητα. Πιο αναλυτικά η ημέρα που παρατηρήθηκε η μέγιστη κινητικότητα ήταν στις 16/8/2020 όπου συμπίπτει με τις διακοπές του Δεκαπενταύγουστου ,επίσης εκείνη την ημέρα δεν είχε εφαρμοστεί το δεύτερο lockdown. Αντίστοιχα η ελάχιστη κινητικότητα παρατηρήθηκε στις 25/12/2020 όπου συμπίπτει με την αργία των Χριστουγέννων, εκείνη την περίοδο είχε εφαρμοσθεί το δεύτερο lockdown. Η ελάχιστη διαμονή στο σπίτι παρατηρήθηκε τις ημέρες 19/7/2020,26/7/2020 και 16/8/2020 ,όλες οι ημέρες αποτελούν Κυριακές και η τρίτη ημέρα είναι μέσα στο διάστημα διακοπών του Δεκαπενταύγουστου,επίσης η ημέρα με την μέγιστη διαμονή στο σπίτι συμπίπτει με την ημέρα ελάχιστης κινητικότητας, το οποίο είναι αναμενόμενο.

Σχετικά με τον επιπολασμό, από τον Πίνακα 3.1 παρατηρείται ελάχιστη τιμή 0.14 που εκφράζεται ως 0.19 κρούσματα ανά 1.000.000 πληθυσμού,η μέγιστη τιμή που έλαβε ο επιπολασμός στο διάστημα μελέτης είναι 414 , δηλαδή 414 άτομα μολύνθηκαν απο τον Covid-19 ανα 1.000.000 πληθυσμού.Το 25% του δείγματος παρουσίασε επιπολασμό μικρότερο απο 12, δηλαδή 12 άτομα ανα 1.000.000 πληθυσμού.Επίσης το 50% του δείγματος παρουσιάζει επιπολασμό μεγαλύτερο από 53.9 ανα 1.000.000 ανθρώπους.

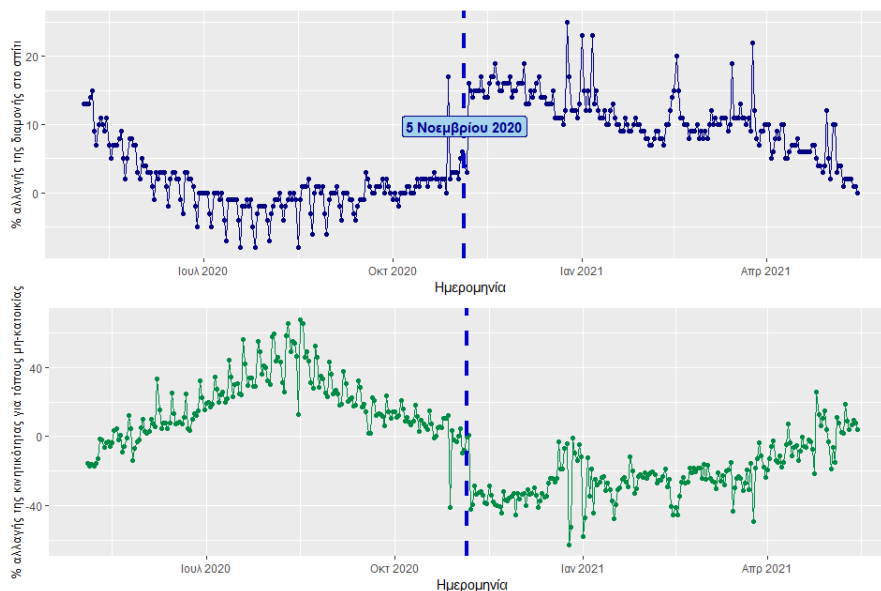
Στο Σχήμα 3.1 παρουσιάζεται η εξέλιξη του επιπολασμού στο διάστημα μελέτης. Η κάθετη διακεκομμένη γραμμή αποτελεί την ημέρα έναρξης της παρέμβασης, του δεύτερου lockdown, δηλαδή χωρίζει το χρονικό διάστημα σε δυο μέρη. Απο το Σχήμα 3.1 παρατηρείται ανοδική τάση των κρουσμάτων μέχρι την παρέμβαση ,έπειτα στο πρώτο τρίμηνο της εφαρμογής του δευτέρου lockdown (Νοέμβριος 2020-Ιανουάριος 2021) ο επιπολασμός και άρα τα κρούσματα μειώνονται. Μετά τον Ιανουάριο 2021 παρατηρείται ανοδική τάση του επιπολασμού και τον Απρίλιο 2021 ο επιπολασμός μεγιστοποιείται. Ο μονόπλευρος έλεγχος t-test για ανεξάρτητα δείγματα με άνισες διακυμάνσεις έδειξε ότι η μέση τιμή του επιπολασμού μετά την παρέμβαση ( $\mu=165.17$ ) είναι σημαντικά μεγαλύτερη ( $p<.001$ ) από εκείνη πριν την παρέμβαση ( $\mu=23$ ). Στο Σχήμα 3.1 στην εικόνα Β παρουσιάζονται τα

error-bars για τον επιπολασμό για κάθε χρονικό διάστημα. Παρατηρείται ότι πριν την παρέμβαση ο επιπολασμός παρουσίαζε μικρότερη μεταβλητότητα συγκριτικά με το δεύτερο lockdown, επίσης είναι εμφανής η διαφορά στις μέσες τιμές στα δυο χρονικά διαστήματα μελέτης αφού τα error-bars δεν επικαλύπτονται.



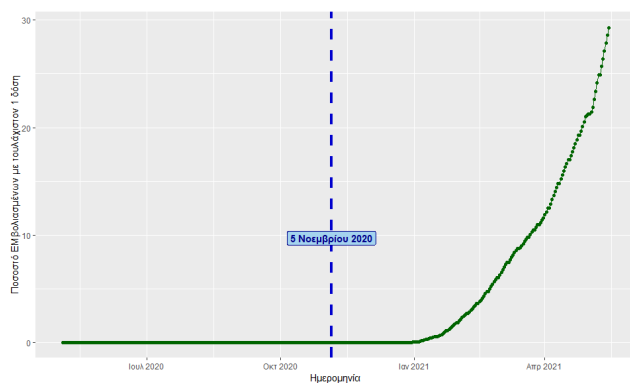
Σχήμα 3.1: A:Εξέλιξη επιπολασμού κατά τη διάρκεια περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown. B: SD Error-Bars για τον επιπολασμό για κάθε χρονικό διάστημα

Ένας παράγοντας που επηρέασε την αποτελεσματικότητα του δεύτερου lockdown και τον αριθμό των κρουσμάτων και κατ'επέκταση την εξάπλωση του Covid-19 είναι οι μετακινήσεις. Στο Σχήμα 3.2 παρουσιάζεται η τάση κινητικότητας πριν και μετά την παρέμβαση. Παρατηρείται ότι πριν την παρέμβαση οι άνθρωποι περνούσαν λιγότερο χρόνο στην κατοικία τους ενώ μετά την παρέμβαση ο χρόνος διαμονής στο σπίτι αυξήθηκε αλλά παρουσιάζει καθοδική τάση, κατά τη διάρκεια της περιόδου μελέτης. Αντίθετα, παρατηρείται ανοδική τάση στο ποσοστό αλλαγής κινητικότητας. Επίσης, μετά την εφαρμογή της παρέμβασης το ποσοστό αλλαγής κινητικότητας για τόπους εκτός κατοικίας είναι αρνητικό εκτός από κάποιες συγκεκριμένες μέρες, δηλαδή η τάση κινητικότητας ήταν αρνητική συγκριτικά με το σημείο αναφοράς. Ο μονόπλευρος έλεγχος t-test με άνισες διακυμάνσεις έδειξε ότι η μέση τιμή του ποσοστού αλλαγής στις μετακινήσεις είναι σημαντικά μικρότερη ( $p < .001$ ) κατά το δεύτερο lockdown ( $\mu = -21.55$ ) συγκριτικά με το διάστημα πριν την εφαρμογή του ( $\mu = 17.63$ ).



Σχήμα 3.2: Ποσοστό αλλαγής στην διαμονή στο σπίτι και στην κινητικότητα συγκριτικά με το σημείο αναφοράς. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown.

Κατά την διάρκεια του δευτέρου lockdown ξεκίνησε ο μαζικός εμβολιασμός, το μέσο ποσοστό εμβολιασμένων κατα το δεύτερο lockdown ήταν 9.17% , δηλαδή περίπου 10 άτομα στα 100 είχαν λάβει τουλάχιστον την πρώτη δόση εμβολίου.

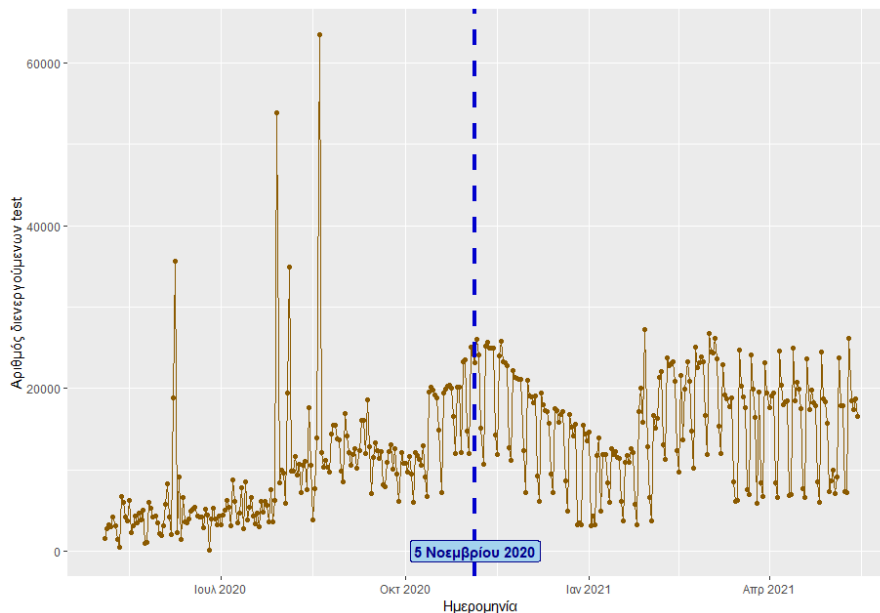


Σχήμα 3.3: Εξέλιξη εμβολιαστικής κάλυψης κατά την διάρκεια της περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown.

Συμπληρωματικά με την εξέλιξη του επιπολασμού και των μετακινήσεων, παρουσιάζεται και η εξέλιξη της εμβολιαστικής κάλυψης στην Ελλάδα. Από τα Διαγράμματα 3.1 παρατηρείται ότι πριν την έναρξη του εμβολιασμού στις 28/12/2020 τα κρούσματα παρουσίασαν καθοδική τάση. Απο το Σχήμα 3.3 παρατηρείται ότι οι εμβολιασμοί συνεχώς αυξάνονταν μέχρι το τέλος το δεύτερου lockdown, γεγονός που να έχει επιδράσει στην αποτελεσματικότητα του lockdown.

Στον περιορισμό της εξάπλωσης του Covid-19 σημαντικό παράγοντα αποτελούσε ο ημερήσιος αριθμός test που διενεργούνταν. Στο διάστημα μελέτης ο μέσος αριθμός ημε-

ρήσιων test ήταν 12825 ,το 50% των ημερών τα test ήταν λιγότερα απο 11864 και στο 25% των ημερών περισσότερα απο 18638. Ο επικρατών ημερήσιος αριθμός test ήταν 4882. Από το Σχήμα 3.4 παρατηρείται ανοδική τάση των test πριν την εφαρμογή της παρέμβασης , με ορισμένες ακραίες τιμές.Μετά την εφαρμογή της παρέμβασης παρατηρείται καθοδική τάση μέχρι τον Ιανουάριο 2021 και έπειτα αύξηση του ημερήσιου αριθμού test χωρίς τάση.

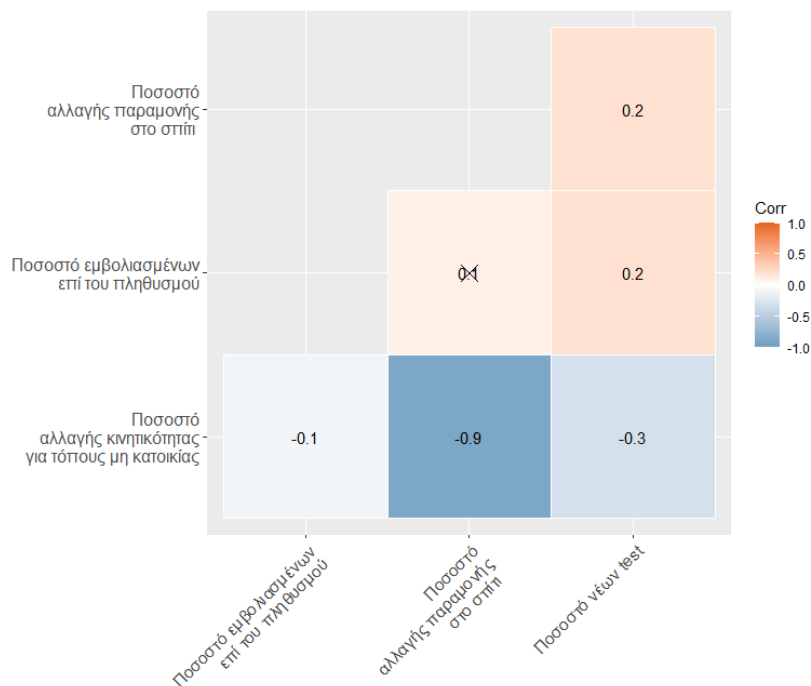


Σχήμα 3.4: Πλήθος Διενεργούμενων test κατά την διάρκεια της περιόδου μελέτης. Η κάθετη διακεκομμένη γραμμή αντιστοιχεί στην ημερομηνία έναρξης του 2ου lockdown.

### 3.2 Στατιστική Ανάλυση μέσω Τμηματικής Παλινδρόμησης (Segmented Regression)

Στη παρούσα εργασία η εξαρτημένη μεταβλητή θεωρείται ο επιπολασμός ανά ένα εκατομμύριο πληθυσμού. Σύμφωνα με προηγούμενες μελέτες (Navazi et al.,2022) ο επιπολασμός εξαρτάται από το συνολικό ημερήσιο ποσοστό εμβολιασμένων ανθρώπων, τις αλλαγές στην κινητικότητα, το ημερήσιο αριθμό test, μεταβλητές σχετικές με την παρέμβαση και τις ημέρες που πέρασαν από την αρχή του διαστήματος μελέτης(  $t=1, 4/5/2020$  ).

Αρχικά διερευνήθηκαν οι συσχετίσεις μεταξύ των παραπάνω μεταβλητών. Ο συντελεστής συσχέτισης του Pearson μεταξύ των μεταβλητών κινητικότητας προέκυψε ίσος με  $-0.92$ . Έτσι διαπιστώνεται ισχυρή αρνητική γραμμική συσχέτιση, επομένως η εισαγωγή και των δύο μεταβλητών στο μοντέλο θα προκαλέσει πρόβλημα πολυσυγγραμμικότητας. Οι συντελεστές συσχέτισης παρουσιάζονται στο Σχήμα 3.5. Το ποσοστό αλλαγής παραμονής στο σπίτι και το ποσοστό εμβολιασμένων επί του πληθυσμού δεν έχουν στατιστικά σημαντική συσχέτιση.



Σχήμα 3.5: Διάγραμμα συσχέτισης συντελεστή Pearson μεταξύ υποψήφιων ανεξάρτητων μεταβλητών.

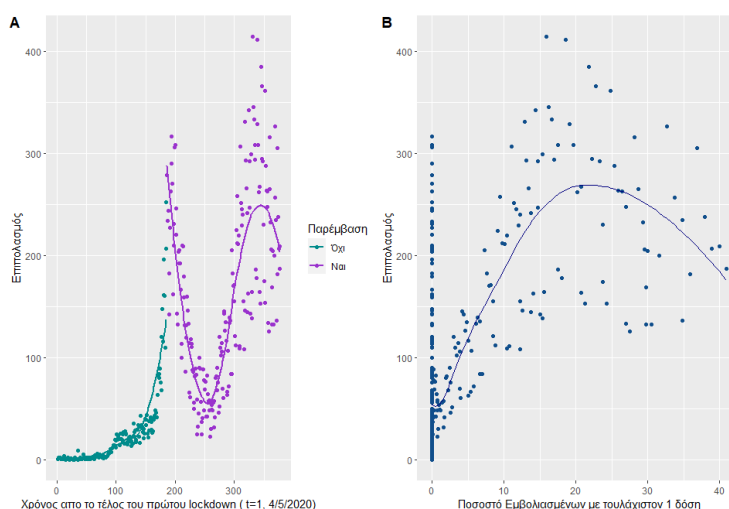
Στον Πίνακα 3.2 συνοψίζονται οι υποψήφιες επεξηγηματικές μεταβλητές. Σχετικά με την μεταβλητή  $TSI_t$  η τιμή 197 δεν αφορά της ημέρα έναρξης του δευτέρου lockdown(5 Νοεμβρίου 2020) αλλά την ημέρα 17 Νοεμβρίου 2020. Σύμφωνα με τους Li et al(2020)

το 95ο εκατοστημόριο της κατανομής της περιόδου επώασης είναι οι 12.5 ημέρες, ενώ σύμφωνα με τους Lauer et al(2020) το 97.5% των ανθρώπων εμφανίζουν συμπτώματα 11.5(CI, 8.2 to 15.6) ημέρες μετά την έκθεση τους στον ιό. Έτσι, οι 12 μέρες αποτελούν τον ελάχιστο χρόνο προκειμένου να δούμε την επίδραση του δεύτερου lockdown στον επιπολασμό. Άρα η 17 Νοέμβρη 2020 ορίζεται ως η ημέρα που θα αρχίσει να φαίνεται η αποτελεσματικότητα της παρέμβασης.

Εξαρτημένη Μεταβλητή	
Prevalence <sub>t</sub>	Επιπολασμός νέων περιπτώσεων ανά ένα εκατομμύριο του πληθυσμού την χρονική στιγμή t
Ανεξάρτητες Μεταβλητές	
t	Ημέρες, ξεκινάει από το τέλος του πρώτου lockdown στην Ελλάδα (t=1, 377)
TVP <sub>t</sub>	Συνολικό Ποσοστό Ανθρώπων που έλαβαν τουλάχιστον 1 δόση την χρονική στιγμή t.
RC <sub>t</sub>	Ποσοστό Αλλαγής Διαμονής στο σπίτι την χρονική στιγμή t
NT <sub>t</sub>	Ποσοστό ημερήσιων test στο πληθυσμό την χρονική στιγμή t
TSl <sub>t</sub>	Αριθμός ημερών από την ημέρα που εφαρμόστηκε η παρέμβαση TSl <sub>t</sub> =0 πριν την εφαρμογή της παρέμβασης TSl <sub>t</sub> =t-197 μετά την εφαρμογή της παρέμβασης
INT <sub>t</sub>	Παρέμβαση INT=0 πριν την εφαρμογή της παρέμβασης INT=1 μετά την εφαρμογή της παρέμβασης
TVP <sub>t</sub> <sup>2</sup>	Μεταβλητή TVP υψωμένη στο τετράγωνο
t <sup>2</sup>	Μεταβλητή t υψωμένη στο τετράγωνο
VC <sub>t</sub>	VC=0 πριν την έναρξη των εμβολιασμών VC=1 μετά την έναρξη των εμβολιασμών.

Πίνακας 3.2: Υποψήφιες επεξηγηματικές μεταβλητές στο μοντέλο τμηματικής παλινδρόμησης με εξαρτημένη μεταβλητή τον επιπολασμό της νόσου

Σε ότι αφορά στους πολυωνυμικούς όρους TVP<sub>t</sub><sup>2</sup> και t<sup>2</sup> που περιλαμβάνονται στον Πίνακα 3.2 θα πρέπει να σημειωθεί ότι κρίνονται υποψήφιες επεξηγηματικές μεταβλητές εξαιτίας της μη γραμμικής σχέσης τους με τον επιπολασμό της νόσου (Σχήμα 3.6). Οι καμπύλες στο παραπάνω σχήμα κατασκευάστηκαν μέσω της μεθόδου Loess και μέσω αυτών είναι εμφανής η ύπαρξη μη γραμμικής σχέσης μεταξύ των μεταβλητών.



Σχήμα 3.6: (Α)Σχέση μεταξύ επιπολασμού και χρόνου, (Β)Σχέση μεταξύ επιπολασμού και ποσοστού εμβολιασμένων.

Μέσω της μεθόδου τμηματικής παλινδρόμησης θα εφαρμοσθούν δυο μοντέλα διακοπόμενων χρονοσειρών . Για το πρώτο μοντέλο θα χρησιμοποιηθεί το πλήρες σύνολο δεδομένων που έχει αναφερθεί, ενώ για το δεύτερο μοντέλο το υπόσυνολο όπου δεν υπάρχει η παρέμβαση, δηλαδή δεδομένα μέχρι την ημερομηνία 10/5/2020. Πιο αναλυτικά στο πρώτο μοντέλο θα περιέχονται όλες οι μεταβλητές του Πίνακα 3.2.

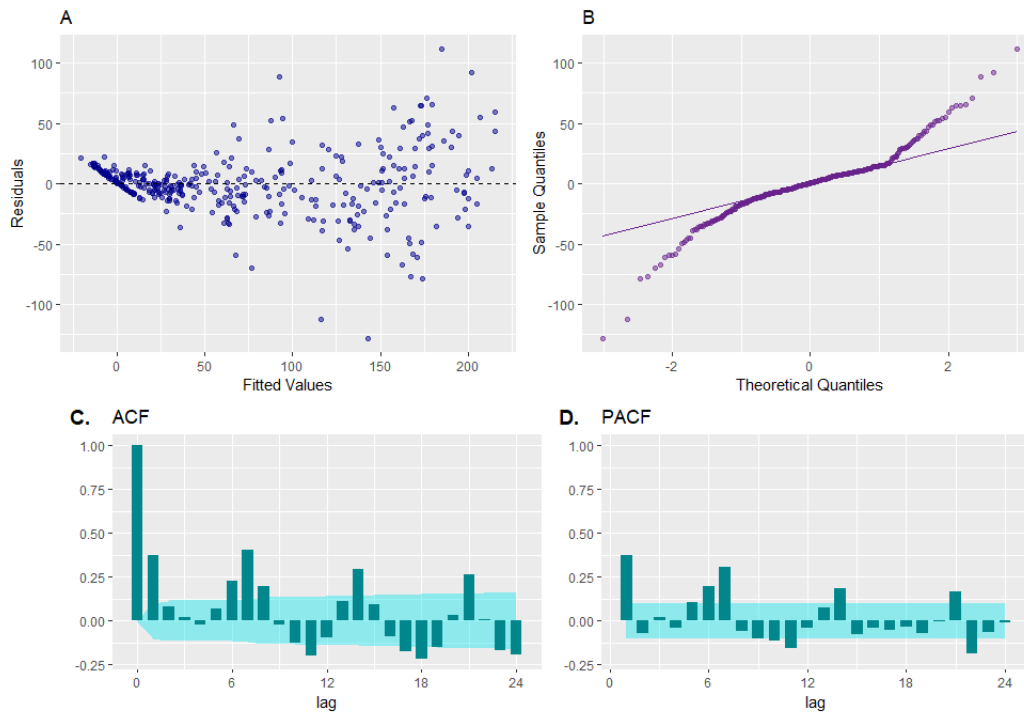
### 3.2.1 Μοντέλο με εφαρμογή παρέμβασης

$$Prevalence_t \sim \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 TVP_t + \beta_4 TVP_t^2 + \beta_5 RC_t + \beta_6 NT_t + \beta_7 TSI_t + \beta_8 INT_t + \beta_9 VC_t + \varepsilon_t \quad (3.1)$$

Η αποτελεσματικότητα του δεύτερου lockdown θα αξιολογηθεί Από τους συντελεστές των μεταβλητών  $TSI_t$  και  $INT_t$ , δηλαδή για να κριθεί ως αποτελεσματική η παρέμβαση θα πρέπει οι συντελεστές  $\beta_7$  και  $\beta_8$  να είναι στατιστικά σημαντικοί και μικρότεροι του μηδενός. Για να αξιολογηθούν οι συντελεστές είναι σημαντικό το μοντέλο να πληρεί τις προϋποθέσεις :

1. Κανονικότητα Καταλοίπων
2. Ανεξαρτησία Καταλοίπων
3. Ομοσκεδαστικότητα Καταλοίπων

Η ομοσκεδαστικότητα των καταλοίπων διερευνήθηκε μέσω του ελέγχου White όπου απορρίφθηκε η υπόθεση της ομοσκεδαστικότητας ( $p\text{-value} < .05$ ). Επίσης η ετεροσκεδαστικότητα των καταλοίπων είναι ορατή και στη εικόνα A του Σχήματος 3.7 , όπου το άπλωμα των καταλοίπων γύρω από τον μέσο δεν είναι σταθερό. Μέσω του ελέγχου Jarque-Berra και Shapiro-Wilk απορρίπτεται η κανονικότητα των καταλοίπων ( $p\text{-value} < .05$ ), αυτό διαπιστώνεται και από την εικόνα B του Σχήματος 3.7 όπου παρατηρούνται μεγάλες αποκλίσεις από την διακεκομμένη γραμμή. Η υπόθεση ανεξαρτησίας των καταλοίπων απορρίφθηκε μέσω του ελέγχου Durbin Watson , όπου και μέσω του διαγράμματος αυτοσυσχέτισης της εικόνας C παρατηρείται ότι σε  $lag=2,6,7,8,10,11,13,14,15,17,18,19,21,23,24,25$  υπάρχει σημαντική αυτοσυσχέτιση των καταλοίπων.



Σχήμα 3.7: Α:Κατάλοιπα έναντι προσαρμοσμένων Τιμών , Β:QQ-Διάγραμμα των καταλοίπων ,  
 C: Διάγραμμα Αυτοσυσχέτισης των καταλοίπων,  
 D: Διάγραμμα Μερικής Αυτοσυσχέτισης των καταλοίπων

Επομένως αφού διαγνώσκει η ύπαρξη ετεροσκεδαστικότητας και αυτοσυσχέτισης των καταλοίπων για την εκτίμηση των συντελεστών  $\beta_i$  θα γίνει χρήση της μεθόδου των Γενικευμένων Ελαχίστων Τετραγώνων (GLS). Τα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης της εικόνας C και D αντίστοιχα του Σχήματος 3.7 παρουσιάζουν ημιτονοειδή μορφή επομένως για την μοντελοποίηση της αυτοσυσχέτισης των καταλοίπων θα χρησιμοποιηθεί ένα αυτοπαλίνδρομο υπόδειγμα κινητών μέσω (ARMA(p,q)). Από το διάγραμμα αυτοσυσχέτισης φαίνεται ότι σε lag=1,6,7,8,11,14,18 υπάρχουν σημαντικά spikes ,από το διάγραμμα μερικής αυτοσυσχέτισης σημαντικά spikes φαίνονται σε lag=6,7,11,14. Άρα το τελικό ARMA(p,q) υπόδειγμα θα αποφασισθεί αφού εφαρμοσθούν υποδείγματα με τους παραπάνω όρους και πραγματοποιηθούν οι κατάλληλες συγκρίσεις και έλεγχοι.

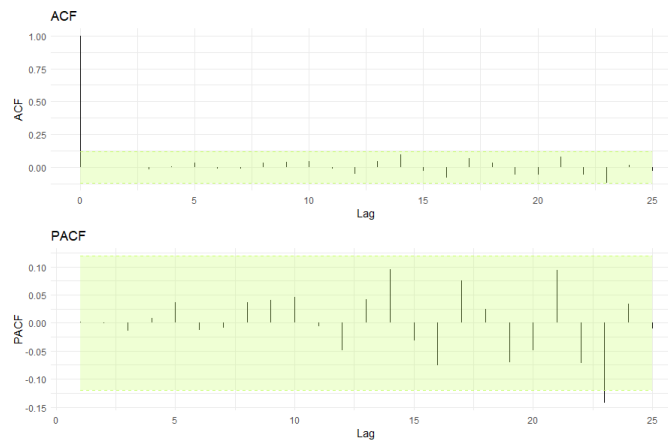
Στον Πίνακα 3.3 συνοψίζονται οι τιμές του κριτηρίου AIC για μοντέλα με αυτοπαλίνδρομο σφάλματα κινητών μέσω τάξης p και q. Παρατηρείται ότι αύξηση της τάξης p και q προκαλεί μείωση του κριτηρίου AIC,



ρ όροι	q όροι			
	1	2	3	4
1	3716.37	3713.49	3718.79	3720.54
2	3715.91	3717.83	3715.44	3699.86
3	3717.79	3716.51	3668.53	3650.45
4	3719.12	3707.00	3670.35	3635.86
5	3692.48	3630.17	3617.82	3664.88
6	3672.18	3613.26	3634.09	3621.48
7	3623.19	3623.98	3603.76	3596.15
8	3598.50	3606.09	3600.67	3584.48
9	3600.07	3600.21	3599.90	3584.14

Πίνακας 3.3: Τιμή κριτηρίου AIC ανα τάξη ρ και q αυτοπαλίνδρομων μοντέλων κινητών μέσω για τα κατάλοιπα το μοντέλου 3.1

Θα χρησιμοποιηθεί το υπόδειγμα ARMA(9,4) για την μοντελοποίηση της αυτοσυσχέτισης των καταλοίπων. Προκειμένου να ελεγχθεί η ύπαρξη αυτοσυσχέτισης των καταλοίπων κατασκευάστηκαν τα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των τυποποιημένων καταλοίπων όπου φαίνονται στο Σχήμα.3.8.

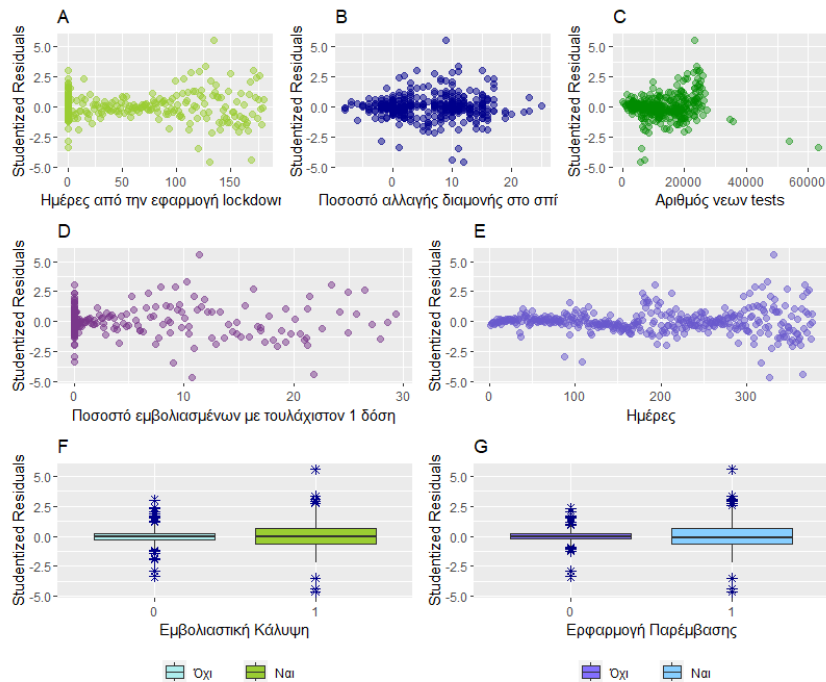


Σχήμα 3.8: Διαγράμματα Αυτοσυσχέτισης και Μερικής αυτοσυσχέτισης τυποποιημένων καταλοίπων

Φαίνεται ότι δεν υπάρχουν κάθετες γραμμές που ξεπερνούν τα όρια της σκιασμένης περιοχής, όπου αποτελεί το 95% ΔΕ. Επομένως τα κατάλοιπα την χρονική στιγμή t μοντελοποιούνται ως εξής:

$$\varepsilon_t = \sum_{p=1}^9 \phi_p \varepsilon_{t-p} + \sum_{q=1}^4 \theta_q e_{t-q} \quad (3.2)$$

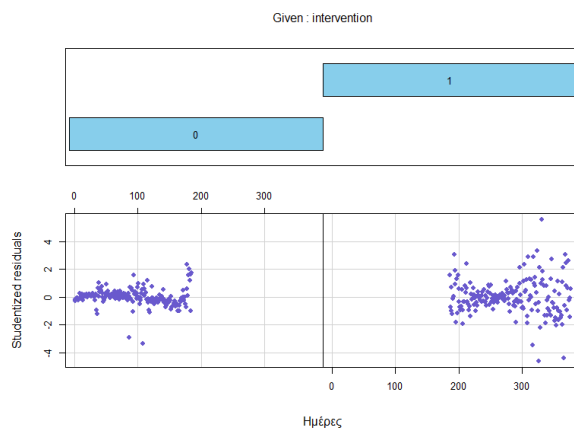
Επόμενο βήμα είναι η αντιμετώπιση της ετεροσκεδαστικότητας των καταλοίπων. Για το σκοπό αυτό θα πρέπει αρχικά να εντοπισθεί η πηγή της διακύμανσης των καταλοίπων. Στο Σχήμα 3.9 παρουσιάζονται τα διαγράμματα των τυποποιημένων καταλοίπων συναρτήσει κάθε ανεξάρτητης μεταβλητής.



Σχήμα 3.9: Τυποποιημένα Κατάλοιπα έναντι των ανεξάρτητων μεταβλητών

Ξεκινώντας από την **Εικόνα Α** παρατηρείται ότι καθώς οι ημέρες αυξάνονται από την αρχή της παρέμβασης υπάρχει μεγαλύτερη μεταβλητότητα, κάτι αντίστοιχο παρατηρείται και στην **Εικόνα Ε** όπου όσο ο αριθμός ημερών αυξάνεται τόσο πιο μεγάλη μεταβλητότητα υπάρχει. Ακόμα από τα διαγράμματα πλαισίου απολήξεων (boxplots) των **Εικόνων F και G** παρατηρείται ότι την περίοδο όπου υπήρχε το lockdown και την περίοδο που είχε ξεκινήσει η εμβολιαστική κάλυψη, όπου συμπίπτει με την πρώτη, υπήρχε μεγαλύτερη μεταβλητότητα. Επίσης από την **Εικόνα Β** παρατηρείται η μεταβλητότητα των τυποποιημένων καταλοίπων να αυξάνεται και μετά μειώνεται καθώς αυξάνεται το ποσοστό αλλαγής της παραμονής στο σπίτι.

Αντίστοιχη συμπεριφορά παρουσιάζουν τα τυποποιημένα κατάλοιπα σε σχέση με τον αριθμό των test που διενεργούνται. Παρατηρώντας το Σχήμα 3.10 τα τυποποιημένα κατάλοιπα ανά επίπεδο της μεταβλητής "Εφαρμογή Παρέμβασης (INT)" είναι ανάλογα της μεταβλητής των ημερών υψωμένη σε μία δύναμη.

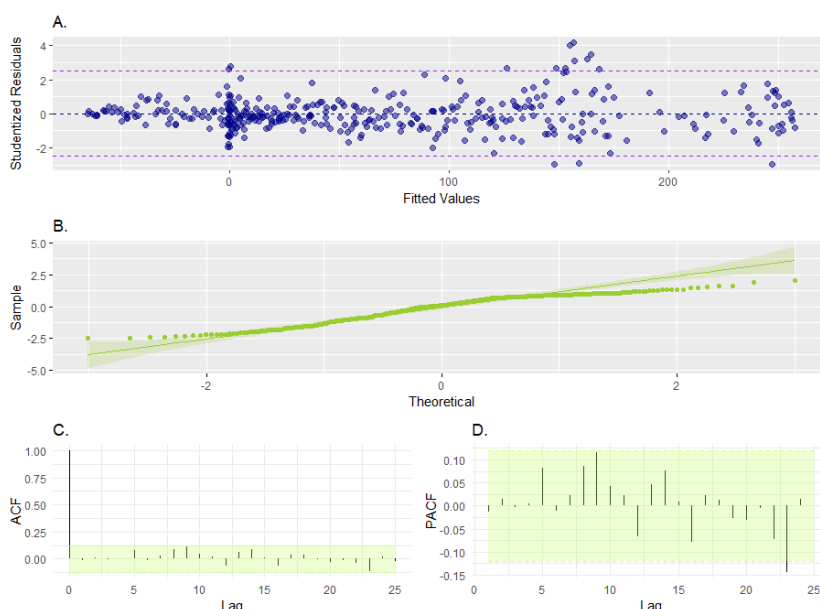


Σχήμα 3.10: Τυποποιημένα Κατάλοιπα έναντι του χρόνου ανά επίπεδο Εφαρμογής Παρέμβασης

Η παραπάνω πληροφορία σχετικά με την μεταβλητότητα των καταλοίπων μπορεί να ενσωματωθεί στο αρχικό μοντέλο. Πιο αναλυτικά πλέον τα κατάλοιπα δεν ακολουθούν την κανονική κατανομή με διακύμανση  $\sigma^2$ , αλλά η διακύμανση μεταβάλλεται καθώς περνούν οι μέρες και είναι διαφορετική για κάθε τμήμα ,δηλαδή πριν και μετά την εφαρμογή της παρέμβασης. Δηλαδή πλέον η διακύμανση των καταλοίπων είναι της μορφής:

$$Var(\varepsilon_t) = \sigma^2 * t_j^{2\delta_j} \quad (3.3)$$

Όπου  $j:1,2$  τα επίπεδα της κατηγορικής μεταβλητής INT,  $t$ :Μέτρηση χρόνου από την λήξη του πρώτου lockdown στην Ελλάδα,  $1, \dots, 377$  . Η παράμετρος  $\delta_j$  εκτιμάται για τις δυο χρονικές περιόδους, πριν και μετά την παρέμβαση, οι τιμές  $\delta_1$  και  $\delta_2$  εκτιμώνται ίσες με 0.94 και 1.07 αντίστοιχα.



Σχήμα 3.11: A: Τυποποιημένα κατάλοιπα έναντι προσαρμοσμένων τιμών του μοντέλου 3.1 με εφαρμογή ARMA(9,4) και διακύμανση της μορφής 3.3  
 B: QQ-Διάγραμμα τυποποιημένων καταλοίπων του μοντέλου 3.1 με εφαρμογή ARMA(9,4) και διακύμανση της μορφής 3.3  
 C: Διάγραμμα αυτοσυσχέτισης τυποποιημένων καταλοίπων  
 D: Διάγραμμα μερικής αυτοσυσχέτισης τυποποιημένων καταλοίπων

Εφαρμόζοντας την παραπάνω δομή διακύμανσης στο μοντέλο με τα αυτοπαλίνδρομο μοντέλο κινητών μέσων (ARMA(9,4)), παρατηρείται ότι τα τυποποιημένα κατάλοιπα εμφανίζουν σταθερό άπλωμα γύρω από το μέσο (μπλε διακεκομμένη γραμμή στην **Εικόνα A**) του Σχήματος 3.11. Επίσης, από την εικόνα B του Σχήματος 3.11 δεν παρατηρείται μεγάλη απόκλιση από την γραμμή αναφοράς, εκτός από τις ουρές της κατανομής που τα σημεία αποκλίνουν από την γραμμή και βρίσκονται εκτός των ορίων του διαστήματος εμπιστοσύνης. Επομένως δεν υπάρχουν σημαντικές ενδείξεις ότι τα τυποποιημένα κατάλοιπα απο-

κλίνουν απο την κανονική κατανομή. Παρατηρώντας τις Εικόνες **C και D** του Σχήματος 3.11 είναι εμφανές ότι τα τυποποιημένα κατάλοιπα δεν εμφανίζουν σημαντική αυτοσυσχέτιση.

Τα αποτελέσματα του μοντέλου με τις διορθώσεις για αυτοσυσχέτιση και ετεροσκεδαστικότητα φαίνονται στον Πίνακα 3.4. Η αποτελεσματικότητα του lockdown θα εξετασθεί απο τους συντελεστές  $\beta_7$  και  $\beta_8$ , δηλαδή θα πρέπει να ελεγχθεί η σημαντικότητα τους καθώς και αν προκάλεσαν κάποια μείωση στην τάση και το επίπεδο.

Συντελεστής	Μεταβλητή	Εκτίμηση	t-value	p-value
$\beta_0$	Σταθερά	-0.652	-0.610	0.542
$\beta_1$	time	-0.169	-0.642	0.521
$\beta_2$	time <sup>2</sup>	0.006	3.196	0.001
$\beta_3$	TVP <sub>t</sub>	44.979	2.844	0.005
$\beta_4$	TVP <sub>t</sub> <sup>2</sup>	-0.649	-2.525	0.012
$\beta_5$	RC <sub>t</sub>	-0.078	0.081	0.336
$\beta_6$	NT <sub>t</sub>	11.760	2.624	0.009
$\beta_7$	TSI <sub>t</sub>	-7.334	-3.815	0.001
$\beta_8$	INT <sub>t</sub>	-12.846	-0.311	0.756
$\beta_9$	VC <sub>t</sub>	35.580	1.488	0.137
Residual standard Error		0.317 (Βαθμοί Ελευθερίας =367)		

Πίνακας 3.4: Εκτιμώμενοι συντελεστές του μοντέλου τμηματικής παλινδρόμησης με εξαρτημένη μεταβλητή τον επιπολασμό της νόσου.

Πιο αναλυτικά ο συντελεστής  $\beta_0$  εκτιμάει το επίπεδο του επιπολασμού την χρονική στιγμή μηδέν. Ο συντελεστής  $\beta_1$  εκτιμάει την ημερήσια αλλαγή στην τάση του επιπολασμού πριν την παρέμβαση (baseline trend). Η αλλαγή στο επίπεδο του επιπολασμού αμέσως μετά την έναρξη του δευτέρου lockdown εκτιμάται απο τον συντελεστή  $\beta_8$ . Τέλος ο συντελεστής  $\beta_7$  εκτιμά την αλλαγή στην τάση του επιπολασμού μετά την παρέμβαση.

Η σημαντικότητα των συντελεστών θα εξετασθεί βάσει του παρακάτω ελέγχου :

$$H_0 : \beta_i = 0 (p - value > .05) \text{ vs } H_1 : \beta_i \neq 0 (p - value < .05)$$

Απο τον Πίνακα 3.4 φαίνεται ότι πριν την έναρξη του δεύτερου lockdown δεν υπήρχε σημαντική ημερήσια αλλαγή στην τάση του επιπολασμού (p=.521). Το δεύτερο lockdown έφερε σημαντική αλλαγή στην τάση του επιπολασμού (p<.001) με ημερήσια μείωση περίπου 7 ατόμων ανά 1.000.000 πληθυσμού, ωστόσο δεν είχε άμεση επίδραση στην μείωση του επιπέδου του επιπολασμού (p=.756).

Πολύ σημαντικό είναι ότι η αλλαγή στο ποσοστό διαμονής στο σπίτι δεν φαίνεται να είχε κάποια σημαντική επίδραση στον έλεγχο του επιπολασμού (RCt: p-value=.336). Ο αριθμός των ημερήσιων test είναι σημαντικός παράγοντας στον έλεγχο της πανδημίας,

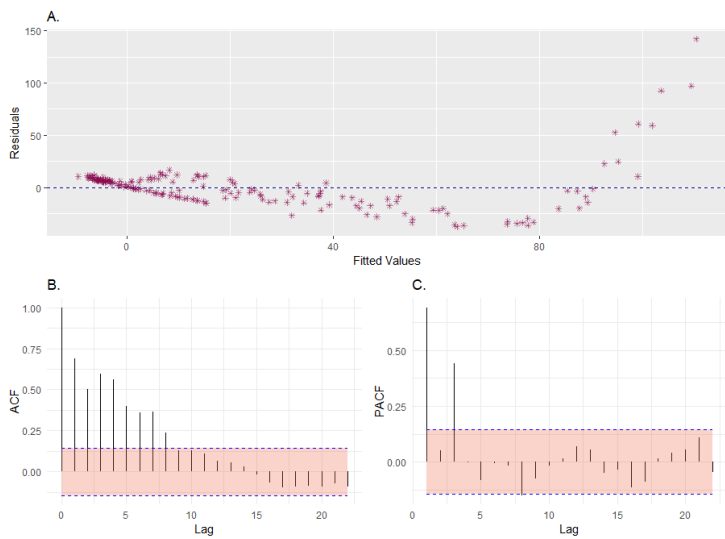
φαίνεται ότι 1% αύξηση στα ημερήσια τεστ προκαλεί ημερήσια αύξηση του επιπολασμού κατά 12 άτομα, όταν οι υπόλοιπες μεταβλητές παραμένουν σταθερές. Η φαρμακευτική παρέμβαση (εμβολιασμός) που έλαβε μέρος στις 28/12/2020 αξιολογείται από τις μεταβλητές TVP & VC. Έτσι βάσει της πρώτης, αύξηση 1% στους ημερήσιους εμβολιασμούς προκαλεί αύξηση του επιπολασμού κατά 45 άτομα ανα 1.000.000. Στην αύξηση αυτή αντικατοπτρίζεται η αυξημένη μεταδοτικότητα του Covid-19 κατά το δεύτερο και τρίτο κύμα. Η μεταβλητή VC από τον Πίνακα 3.4 δεν φαίνεται στατιστικά σημαντική.

### 3.2.2 Counterfactual Σενάριο

Προκειμένου να εκτιμηθεί το υποθετικό σενάριο της εξέλιξης του επιπολασμού αν δεν είχε εφαρμοσθεί ποτέ το δεύτερο lockdown εφαρμόστηκε ένα καινούριο μοντέλο στα δεδομένα πριν την παρέμβαση αφαιρώντας από το μοντέλο τις μεταβλητές που είναι σχετικές με την παρέμβαση. Ωστόσο επειδή οι εμβολιασμοί είχαν ξεκινήσει κατά την διάρκεια του lockdown αφαιρέθηκαν και αυτές οι μεταβλητές. Το καινούριο πλήρες μοντέλο γράφεται ως εξής:

$$Prevalence_t \sim \beta_0 + \beta_1 time + \beta_2 time^2 + \beta_3 RC_t + \beta_4 NT_t + \varepsilon_t, t = 1, \dots, 185 \quad (3.4)$$

Ελέγχοντας τις υποθέσεις των καταλοίπων μέσω των ελέγχων και διαγραμμάτων που εφαρμόστηκαν και στο αρχικό μοντέλο διαπιστώθηκε ότι τα κατάλοιπα παρουσιάζουν ετεροσκεδαστικότητα και αυτοσυσχέτιση. Η ύπαρξη ετεροσκεδαστικότητας είναι εμφανή στην **Εικόνα Α** του Σχήματος 3.12 καθώς τα κατάλοιπα δεν κατανέμονται ομοιόμορφα γύρω από την οριζόντια γραμμή. Παρατηρώντας τις **Εικόνες Β και C** του Σχήματος 3.12 φαίνονται σημαντικά spikes στα διαγράμματα αυτοσυσχέτισης και μερικής αυτοσυσχέτισης των καταλοίπων. Αναλυτικότερα, στην Εικόνα Β του Σχήματος 3.12 παρατηρείται ότι καθώς τα lag αυξάνονται ο συντελεστής συσχέτισης συγκλίνει στο μηδέν, ενώ στην Εικόνα C του Σχήματος 3.12 ο συντελεστής μερικής αυτοσυσχέτισης τείνει απότομα προς το μηδέν μετά το τρίτο lag. Βάσει αυτών, για την διόρθωση της αυτοσυσχέτισης θα εφαρμοσθεί το κατάλληλο αυτοπαλίνδρομο υπόδειγμα. Εφαρμόζοντας το αυτοπαλίνδρομο υπόδειγμα με 9 AR όρους στα κατάλοιπα φαίνεται στις Εικόνες C και D του Σχήματος 3.13 ότι η αυτοσυσχέτιση των καταλοίπων διορθώθηκε. Μάλιστα ο συντελεστής Durbin-Watson για τις χρονικές στιγμές 1 έως 14 προέκυψε 2, όπου αυτό φανερώνει την ανεξαρτησία των τυποποιημένων καταλοίπων.



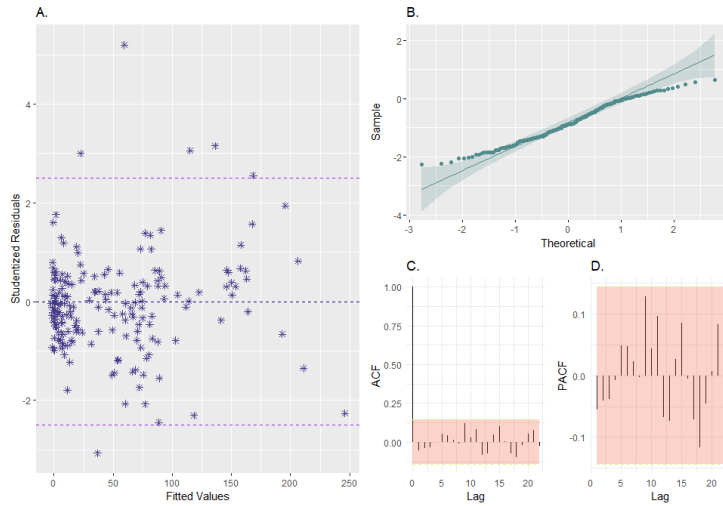
Σχήμα 3.12: Α: Κατάλοιπα έναντι προσαρμοσμένων τιμών

του μοντέλου 3.4, Β: Διάγραμμα αυτοσυσχέτισης καταλοίπων του μοντέλου 3.4, C: Διάγραμμα μερικής αυτοσυσχέτισης καταλοίπων του μοντέλου 3.4

Για την διόρθωση της ετεροσκεδασικότητας η διακύμανση των καταλοίπων τελικά μοντελοποιήθηκε σύμφωνα με την εξαρτημένη μεταβλητή NT (αριθμός ημερήσιων τεστ) ως εξής:

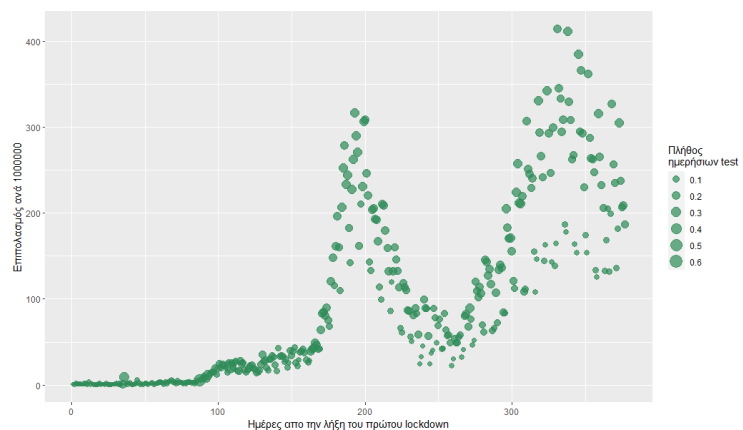
$$Var(\varepsilon_t) = \sigma^2 * (NT_t)^{2\delta} \quad (3.5)$$

όπου η σταθερά  $\delta$  εκτιμήθηκε από το μοντέλο ίση με 1.09. Τα διαγνωστικά διαγράμματα φαίνονται



Σχήμα 3.13: A:Τυποποιημένα Κατάλοιπα έναντι προσαρμοσμένων τιμών του μοντέλου 3.4  
 B: QQ-Διάγραμμα τυποποιημένων καταλοίπων του μοντέλου 3.4 με εφαρμογή AR(9) και διακύμανση 3.5  
 C: Διάγραμμα αυτοσυσχέτισης τυποποιημένων καταλοίπων του μοντέλου 3.4 με εφαρμογή AR(9) και διακύμανση 3.5  
 D: Διάγραμμα μερικής αυτοσυσχέτισης τυποποιημένων καταλοίπων του μοντέλου 3.4 με εφαρμογή AR(9) και διακύμανση 3.5

Απο το Σχήμα 3.14 παρατηρείται ότι καθώς οι ημέρες από το τέλος του πρώτου lockdown (t) αυξάνονται, για τον ίδιο αριθμό test (NT) ο επιπολασμός της νόσου λαμβάνει μεγαλύτερες τιμές. Προκειμένου λοιπόν, να ληφθεί υπόψιν η σχέση μεταξύ των τριών μεταβλητών προστέθηκε στο μοντέλο 3.4 ο όρος της αλληλεπίδρασης μεταξύ της μεταβλητής "t" και "NT". Το τελικό μοντέλο γράφεται ως εξής:



Σχήμα 3.14: Διάγραμμα διασποράς των μεταβλητών Επιπολασμός, Ημέρες από το τέλος του πρώτου lockdown και αριθμό ημερήσιων test.

$$Prevalence_t \sim \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 NT_t + \beta_4 NT_t * t + \beta_5 RC_t + \varepsilon_t, t = 1, \dots, 185 \quad (3.6)$$

Στον Πίνακα 3.5 παρουσιάζονται τα αποτελέσματα του μοντέλου 3.6

Συντελεστής	Μεταβλητή	Εκτίμηση	t-value	p-value
$\beta_0$	Σταθερά	1.83	2.27	0.02
$\beta_1$	time <sup>2</sup>	-0.02	6.69	0.01
$\beta_2$	time	-0.09	-4.32	0.01
$\beta_3$	NT <sub>t</sub>	-102.32	-0.74	0.46
$\beta_4$	NT <sub>t</sub> *t	4.52	2.43	0.02
$\beta_5$	RC <sub>t</sub>	-0.01	-0.19	0.8
<b>Residual Standard Error =369.555</b>		<b>Βαθμοί Ελευθερίας =179</b>		

Πίνακας 3.5: Αποτελέσματα μοντέλου (3.6)

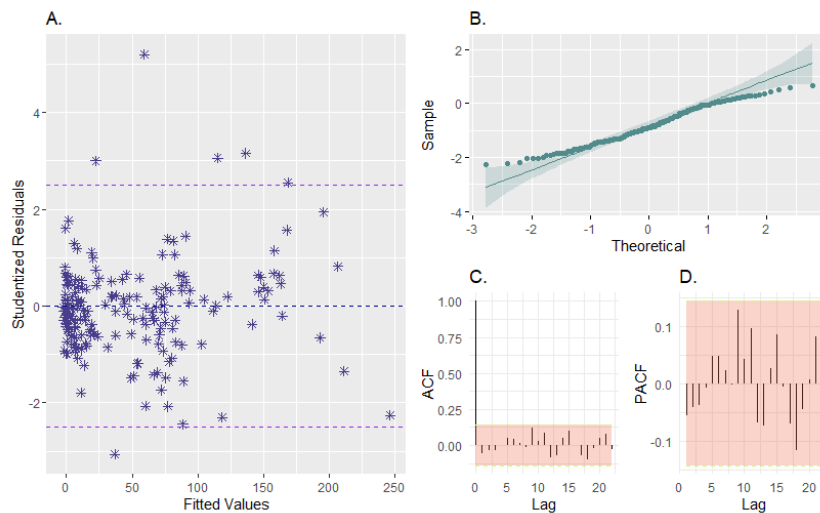
Φαίνεται ότι η μεταβλητή RC (ποσοστό αλλαγής παραμονής στο σπίτι) παραμένει μη στατιστικά σημαντική, αφού αφαιρεθεί το τελικό μοντέλο γράφεται όπως η σχέση 3.6 εξαιρώντας την μεταβλητή RC<sub>t</sub>.

Πιο αναλυτικά το τελικό μοντέλο μέσω του οποίου θα εκτιμηθεί η εξέλιξη του επιπολασμού χωρίς την

εφαρμογή παρέμβασης γράφεται ως εξής:

$$Prevalence_t \sim \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 NT_t + \beta_4 NT_t * t + \varepsilon_t, t = 1, \dots, 185 \quad (3.7)$$

Από το Σχήμα 3.15, που περιλαμβάνει διαγνωστικά διαγράμματα των τυποποιημένων καταλοίπων φαίνεται ότι δεν υπάρχουν σημαντικές ενδείξεις παραβίασης των υποθέσεων καθώς σε όλες τις εικόνες τα περισσότερα σημεία βρίσκονται μέσα στην ζώνη εμπιστοσύνης.



Σχήμα 3.15: A:Τυποποιημένα Κατάλοιπα έναντι προσαρμοσμένων τιμών μοντέλου (3.7)  
 B: Q-Q-plot τυποποιημένων καταλοίπων μοντέλου (3.7)  
 C: Διάγραμμα αυτοσυσχέτισης τυποποιημένων καταλοίπων μοντέλου (3.7)  
 D: Διάγραμμα μερικής αυτοσυσχέτισης τυποποιημένων καταλοίπων μοντέλου (3.7)

Τα αποτελέσματα του τελικού μοντέλου παρουσιάζονται στον Πίνακα 3.6. Από το συντελεστή του όρου της αλληλεπίδρασης φαίνεται ότι κάθε επιπλέον ημέρα αυξάνει την επί-

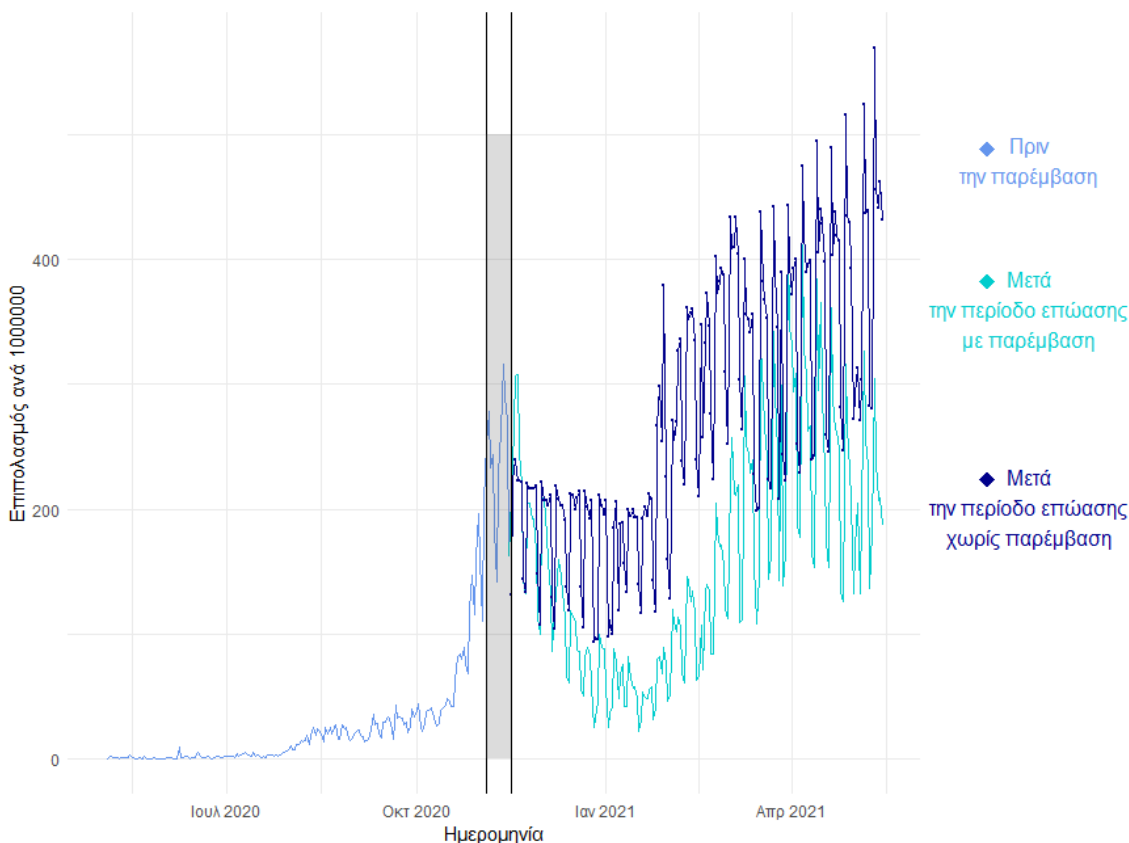


δραση των test που γίνονται στον επιπολασμό κατά 4 άτομα ανά 1.000.000 πληθυσμού.

Συντελεστής	Μεταβλητή	Εκτίμηση	t-value	p-value
b <sub>0</sub>	Σταθερά	1.68	5.93	0.01
b <sub>1</sub>	time <sup>2</sup>	-0.09	11.31	0.01
b <sub>2</sub>	time	-0.09	-6.54	0.01
b <sub>3</sub>	NT <sub>t</sub>	-96.39	-0.74	0.46
b <sub>4</sub>	NT <sub>t</sub> *time	4.51	2.44	0.01
<b>Residual Standard Error= 372.72</b>		<b>Βαθμοί Ελευθερίας =180</b>		

Πίνακας 3.6: Αποτελέσματα τελικού μοντέλου(3.7)

Στη συνέχεια, συγκρίνεται ο επιπολασμός του Covid-19 μετά την έναρξη της παρέμβασης στις 5 Νοεμβρίου 2020, με τον εκτιμώμενο επιπολασμό για την ίδια περίοδο χωρίς την εφαρμογή της παρέμβασης. Ο εκτιμώμενος επιπολασμός χωρίς την εφαρμογή του lockdown φαίνεται στο Σχήμα 3.16. Η σκιασμένη περιοχή αφορά τις 12 ημέρες που όπως προαναφέρθηκε στην αρχή της ανάλυσης είναι η περίοδος η επώασης, δηλαδή ο χρόνος που θα χρειαστεί να εμφανιστούν τα συμπτώματα μετά την έκθεση στον Covid-19.



Σχήμα 3.16: Σύγκριση εξέλιξης επιπολασμού και counterfactual σενάριο

Από το Σχήμα 3.16 παρατηρείται ότι με την εφαρμογή του lockdown αρχικά ο επιπολασμός παρέμεινε αρκετά μικρότερος σε σχέση με τον εκτιμώμενο χωρίς παρέμβαση. Ωστόσο από τις 2 Φεβρουαρίου 2021 φαίνεται ότι ο εκτιμώμενος επιπολασμός προσεγγίζει τον πραγματικό. Επίσης όταν ξεκίνησαν οι εμβολιασμοί φαίνεται ότι το lockdown συνέβαλλε στον περιορισμό της εξάπλωσης του επιπολασμού, το οποίο ήταν απαραίτητο

για την εξασφάλιση ενός σημαντικού ποσοστού μη μολυσμένων ανθρώπων (Lytras and Tsiodras,2021).

Η μέση τιμή του επιπολασμού ανά εκατομμύριο μετά τις 5/11/2020 με την εφαρμογή του lockdown είναι 165 όπου θα ήταν 257 αν δεν είχε εφαρμοσθεί. Άρα το lockdown προκάλεσε κατα μέσο όρο 39.68% ποσοστιαία μείωση στον μέση τιμή του επιπολασμού. Παρακάτω παρουσιάζεται συνοπτικός πίνακας με την ποσοστιαία μεταβολή που προκάλεσε το lockdown στον επιπολασμο.

	Ελάχιστη Τιμή	1ο τεταρτημόριο	Διάμεσος	Μέση Τιμή	3ο τεταρτημόριο	Μέγιστη Τιμή
<b>Με lockdown</b>	22.61	86.17	145.91	165.17	233.84	414.02
<b>Χωρίς lockdown</b>	98.35	202.26	238.90	273.83	372.77	568.98
<b>Μεταβολή</b>	77.78%	57.40%	38.92%	39.68%	37.27%	27.23%

Πίνακας 3.7: Συνοπτικός Πίνακας μεταβολής του επιπολασμού λόγω του lockdown

Από τον Πίνακα 3.7 φαίνεται ότι η εφαρμογή του lockdown προκάλεσε ποσοστιαία μείωση 27.23% στην μέγιστη τιμή του επιπολασμού, η οποία είναι αρκετά μικρότερη απο αυτή της ελάχιστης τιμής.

## Κεφάλαιο 4

### Συμπεράσματα

Η μελέτη αυτή ανέδειξε την χρησιμότητα των μοντέλων ανάλυσης διακοπτόμενων χρονοσειρών για την κατανόηση των δεδομένων Covid-19. Οι παρεμβάσεις, όπως lockdowns, έχουν σημαντική επίδραση στη διασπορά του ιού και η αποτελεσματική ανάλυση πρέπει να λαμβάνει υπόψη τις παρεμβάσεις αυτές. Τα μοντέλα ανάλυσης διακοπτόμενων χρονοσειρών μπορούν να χρησιμοποιηθούν για προβλέψεις και ανάπτυξη στρατηγικών αντιμετώπισης. Αυτό είναι ιδιαίτερα χρήσιμο στον προγραμματισμό πιο αποτελεσματικών δράσεων και στην προσαρμογή μέτρων στις μεταβαλλόμενες συνθήκες της πανδημίας.

Κατέδειξε πως το δεύτερο lockdown είχε θετική επίδραση στον περιορισμό της εξάπλωσης του ιού. Η εφαρμογή του lockdown βοήθησε στη μείωση του αριθμού των κρουσμάτων, αποτρέποντας έτσι την υπερφόρτωση του υγειονομικού συστήματος. Ωστόσο η παραμονή στο σπίτι φάνηκε να μην έπαιξε σημαντικό ρόλο στον περιορισμό της εξάπλωσης, δεδομένου ότι την χρονική περίοδο της μελέτης υπήρχαν μερικές άρσεις των μέτρων αυτό συνέβαλε στο αποτέλεσμα.

Καθώς είναι απαραίτητο να εξισορροπήσουμε την προστασία της δημόσιας υγείας με τις αρνητικές συνέπειες που προκαλεί ένα αυστηρό lockdown και ο εκτιμώμενος επιπολασμός των κρουσμάτων προσέγγισε τον πραγματικό κατά τη λήξη του δεύτερου, τίθεται το ερώτημα αν θα ήταν απαραίτητο η διάρκεια του lockdown να ήταν τόσο εκτεταμένη.

Μία εναλλακτική μέθοδος ανάλυσης διακοπτόμενων χρονοσειρών είναι μέσω των αυτοπαλίνδρομων ολοκληρωμένων υποδειγμάτων κινητών μέσων (ARIMA). Τα πλεονεκτήματα της συγκεκριμένης μεθόδου έναντι της τμηματικής παλινδρόμησης είναι ότι με τα ARIMA υποδείγματα μπορεί να εντοπισθεί και να διαχειρισθεί η εποχικότητα καθώς και μη στάσιμες χρονοσειρές. Ωστόσο στα ARIMA υποδείγματα μπορεί να υπάρξει μεροληψία στις εκτιμήσεις των συντελεστών που είναι σχετικές με την παρέμβαση λόγω αλλαγής του τρόπου μέτρησης της μεταβλητής απόκρισης στην πάροδο του χρόνου. (Li et al., 2021)

# Βιβλιογραφία

- Aitken, A. (1935). IV.—On Least Squares and Linear Combination of Observations. *Proceedings of the Royal Society of Edinburgh*, 55, 42-48. [doi:10.1017/S0370164600014346](https://doi.org/10.1017/S0370164600014346)
- Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted time series regression for the evaluation of public health interventions: a tutorial. *International journal of epidemiology*, 46(1), 348–355. <https://doi.org/10.1093/ije/dyw098>
- Covid-19 community mobility reports - google. <https://www.google.com/covid19/mobility/>
- Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (CoVID-19) from publicly reported confirmed cases: Estimation and application. *Ann Intern Med*. 2020; 172: 577–582. <https://doi.org/10.7326/M20-0504>
- Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N Engl J Med*. 2020; 382: 1199–1207. <https://doi.org/10.1056/NEJMoa2001316>
- Li, L., Cuerden, M. S., Liu, B., Shariff, S., Jain, A. K., amp; Mazumdar, M. (2021). Three statistical approaches for assessment of intervention effects: A Primer for Practitioners. *Risk Management and Healthcare Policy*, Volume 14, 757–770. <https://doi.org/10.2147/rmhp.s275831>

- Linear mixed-effects models using r: A step-by-step approach. (2013). New York etc.: Springer.
- Lytras, T., & Tsiodras, S. (2021). Lockdowns and the COVID-19 pandemic: What is the endgame?. *Scandinavian journal of public health*, 49(1), 37-40.
- Navazi, F., Yuan, Y., & Archer, N. (2022). The effect of the Ontario stay-at-home order on Covid-19 third wave infections including vaccination considerations: An interrupted time series analysis. *PLoS One*, 17(4), e0265549.  
<https://doi.org/10.1371/journal.pone.0265549>
- Zuur, A. F., Ieno, E. N., Walker, N. J., Saveliev, A. A., Smith, G. M. (2008). *Mixed Effects Models and Extensions in Ecology with R*. Springer.  
<https://doi.org/10.1007/978-0-387-87458-6>
- U.S. Department of Health and Human Services. (n.d.). What is prevalence?. National Institute of Mental Health.  
<https://www.nimh.nih.gov/health/statistics/what-is-prevalence>  
<https://www.nimh.nih.gov/health/statistics/what-is-prevalence>
- Woc-Colburn, L., & Godinez, D. (2022). Lockdown as a public health measure. *COVID-19 Pandemic*, 133–136.  
<https://doi.org/10.1016/B978-0-323-82860-4.00013-6>
- Ατσαλάκης, Γ. Σ. (2021, November 17). Το πιο επικίνδυνο κύμα της πανδημίας. ΚΡΗΤΩΝ ΕΠΙΧΕΙΡΕΙΝ.  
<https://cretanbusiness.gr/%CF%84%CE%BF-%CF%80%CE%B9%CE%BF-%CE%B5%CF%80%CE%B9%CE%BA%CE%AF%CE%BD%CE%B4%CF%85%CE%BD%CE%BF-%CE%BA%CF%8D%CE%BC%CE%B1-%CF%84%CE%B7%CF%82-%CF%80%CE%B1%CE%BD%CE%B4%CE%B7%CE%BC%CE%AF%CE%B1%CF%82/>
- Τζαβαλής, Η. (2008). *Οικονομετρία*. ΟΠΑ.