



«ΑΝΑΛΥΣΗ ΛΟΓΟΥ ΚΑΙ ΣΥΝΑΙΣΘΗΜΑΤΟΣ ΤΗΣ FEDERAL OPEN MARKET COMMITTEE ΣΤΗΝ ΠΑΝΔΗΜΙΑ COVID-19 ΜΕ ΤΕΧΝΙΚΕΣ ΤΕΧΝΗΤΗΣ ΝΟΗΜΟΣΥΝΗΣ»

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΤΥΡΛΗ ΑΙΚΑΤΕΡΙΝΗ-ΛΙΛΙΚΑ

ΑΜ: 4200136

ΕΠΙΒΛΕΠΟΥΣΑ ΚΑΘΗΓΗΤΡΙΑ: ΒΕΛΛΑ ΕΥΓΕΝΙΑ

ΙΑΝΟΥΑΡΙΟΣ 2026

« Η ΤΥΡΛΗ ΑΙΚΑΤΕΡΙΝΗ-ΛΙΛΙΚΑ ΒΕΒΑΙΩΝΩ ΟΤΙ ΤΟ ΕΡΓΟ ΠΟΥ ΕΚΠΟΝΗΘΗΚΕ ΚΑΙ ΠΑΡΟΥΣΙΑΖΕΤΑΙ ΣΤΗΝ ΥΠΟΒΑΛΛΟΜΕΝΗ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ ΕΙΝΑΙ ΑΠΟΚΛΕΙΣΤΙΚΑ ΑΤΟΜΙΚΑ ΔΙΚΟ ΜΟΥ. ΌΠΟΙΕΣ ΠΛΗΡΟΦΟΡΙΕΣ ΚΑΙ ΥΛΙΚΟ ΠΟΥ ΠΕΡΙΕΧΟΝΤΑΙ ΕΧΟΥΝ ΑΝΤΛΗΘΕΙ ΑΠΟ ΑΛΛΕΣ ΠΗΓΕΣ, ΕΧΟΥΝ ΚΑΤΑΛΛΗΛΩΣ ΑΝΑΦΕΡΘΕΙ ΣΤΗΝ ΠΑΡΟΥΣΑ ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ. ΕΠΙΠΛΕΟΝ ΤΕΛΩ ΕΝ ΓΝΩΣΕΙ ΟΤΙ ΣΕ ΠΕΡΙΠΤΩΣΗ ΔΙΑΠΙΣΤΩΣΗΣ ΟΤΙ ΔΕΝ ΣΥΝΤΡΕΧΟΥΝ ΟΣΑ ΒΕΒΑΙΩΝΟΝΤΑΙ ΑΠΟ ΜΕΡΟΥΣ ΜΟΥ Η ΕΡΓΑΣΙΑ ΜΟΥ ΘΑ ΜΗΔΕΝΙΣΤΕΙ.»

ΠΕΡΙΛΗΨΗ:

Η παρούσα διπλωματική εργασία διερευνά την επικοινωνία της Επιτροπής Ανοιχτής Αγοράς (Federal Open Market Committee - FOMC) της Federal Reserve των ΗΠΑ κατά την περίοδο 2018–2024, με ιδιαίτερη έμφαση στην ανταπόκριση στην πανδημία COVID-19. Μέσω πολυμεθοδικής προσέγγισης της Επεξεργασίας Φυσικής Γλώσσας (NLP)και τεχνικών Τεχνητής Νοημοσύνης (AI), συνδυάζει ανάλυση συναισθήματος (με το εξειδικευμένο λεξικό Loughran-McDonald και το μοντέλο FinBERT) και θεματική ανάλυση (LDA) για να εξάγει ποσοτικά και ποιοτικά συμπεράσματα τόσο για τον συναισθηματικό τόνο («πώς») όσο και για το περιεχόμενο («τί») των Πρακτικών (Minutes) της FOMC.

Τα αποτελέσματα καταδεικνύουν ότι ο γλωσσικός τόνος των πρακτικών είναι σταθερά προσεκτικός με έντονη αρνητική τροπικότητα σε περιόδους κρίσης (2020, 2022–2023), λειτουργώντας ως ευαίσθητος δείκτης οικονομικής αβεβαιότητας. Η θεματική ανάλυση αποκαλύπτει μια δυναμική μετατόπιση προτεραιοτήτων: από την προ-πανδημική εστίαση στον πληθωρισμό και την ανάπτυξη, προς τη διαχείριση χρηματοπιστωτικών κινδύνων κατά την πανδημία, και στη συνέχεια στην επανεστίαση στον πληθωρισμό και τη σταθεροποίηση της αγοράς εργασίας το 2024. Η έρευνα επιχειρεί επίσης να συσχετίσει τα γλωσσικά ευρήματα με τις πραγματικές πολιτικές παρεμβάσεις, διαπιστώνοντας ότι ο λόγος της FOMC αντανακλά άμεσα τη διαδικασία λήψης νομισματικών αποφάσεων, ενώ έχει περιορισμένη προγνωστική ισχύ ως προς μελλοντικούς οικονομικούς δείκτες.

Η μελέτη συνεισφέρει τόσο στη βελτίωση της εμπειρικής κατανόησης της νομισματικής επικοινωνίας σε περιόδους κρίσης, όσο και στη μεθοδολογική συζήτηση σχετικά με την πολυμεθοδική ανάλυση πολύπλοκων θεσμικών κειμένων, υπογραμμίζοντας τη συμβολή των τεχνικών «κειμένου ως δεδομένων» στην οικονομική έρευνα.

ΠΕΡΙΕΧΟΜΕΝΑ:

ΠΕΡΙΛΗΨΗ: 2

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ-ΠΙΝΑΚΩΝ-ΣΧΗΜΑΤΩΝ: 5

ΛΕΞΙΚΟ ΟΡΩΝ: 6

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ 11

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΚΑΙ ΒΙΒΛΙΟΓΡΑΦΙΚΟ ΠΛΑΙΣΙΟ 13

2.1 ΠΡΟΣΕΓΓΙΣΗ «ΚΕΙΜΕΝΟ ΩΣ ΔΕΔΟΜΕΝΟ» 14

2.2 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (SENTIMENT ANALYSIS) 15

2.2.1 ΛΕΞΙΚΟΛΟΓΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ 15

2.2.2 ΠΡΟΣΕΓΓΙΣΕΙΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ - ΕΠΑΝΑΣΤΑΣΗ ΤΟΥ FINBERT 16

2.3 ΘΕΜΑΤΙΚΗ ΑΝΑΛΥΣΗ (TOPIC MODELING) ΚΑΙ ΜΟΝΤΕΛΟ LATENT DIRICHLET ALLOCATION (LDA) 17

2.4 ΠΡΟΗΓΟΥΜΕΝΕΣ ΕΦΑΡΜΟΓΕΣ ΣΕ FOMC MINUTES 18

2.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΒΙΒΛΙΟΓΡΑΦΙΚΗΣ ΑΝΑΣΚΟΠΗΣΗΣ 19

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ 21

3.1 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ 21

3.2 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΕΙΜΕΝΟΥ 23

3.2.1 ΔΙΑΔΙΚΑΣΙΑ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑΣ 23

3.2.2 ΑΠΟΘΗΚΕΥΣΗ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗ ΔΕΔΟΜΕΝΩΝ 25

3.3 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (SENTIMENT ANALYSIS) 26

3.3.1 ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ 26

3.3.2 ΠΡΟΣΕΓΓΙΣΗ ΜΟΝΤΕΛΟΥ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ 27

3.4 ΘΕΜΑΤΙΚΗ ΑΝΑΛΥΣΗ (LATENT DIRICHLET ALLOCATION-LDA) 28

3.4.1 ΥΛΟΠΟΙΗΣΗ LDA 29

3.4.2 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ 29

3.4.2.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ 29

3.4.2.2 ΔΗΜΙΟΥΡΓΙΑ ΛΕΞΙΛΟΓΙΟΥ 30

3.4.2.3 ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ LDA 30

3.4.3 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ 31
3.4.3.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ 31
3.4.3.2 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΕΙΜΕΝΟΥ ΓΙΑ ΚΑΘΕ ΕΤΟΣ 32
3.4.3.3 ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ LDA 32
3.5 ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΜΕΘΟΔΟΛΟΓΙΩΝ 33
3.6 ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ 34
ΚΕΦΑΛΑΙΟ 4: ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ 35
4.1 ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ 36
4.1.1 ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ 36
4.1.2 ΠΡΟΣΕΓΓΙΣΗ FINBERT 40
4.2 ΑΠΟΤΕΛΕΣΜΑΤΑ ΘΕΜΑΤΙΚΗΣ ΑΝΑΛΥΣΗΣ 44
4.2.1 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ 44
4.2.2 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ 47
4.2.2.1 ΠΡΟ-ΠΑΝΔΗΜΙΚΗ ΠΕΡΙΟΔΟΣ(2018-2019) 47
4.2.2.2 ΠΕΡΙΟΔΟΣ ΠΑΝΔΗΜΙΑΣ COVID-19(2020-2021) 48
4.2.2.3 ΜΕΤΑ-ΠΑΝΔΗΜΙΚΗ ΚΑΙ ΠΛΗΘΩΡΙΣΤΙΚΗ ΠΕΡΙΟΔΟΣ(2022-2023) 50
4.2.2.4 ΠΕΡΙΟΔΟΣ ΣΤΑΘΕΡΟΠΟΙΗΣΗΣ(2024) 51
4.3 ΣΥΝΟΛΙΚΗ ΕΡΜΗΝΕΙΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ 54
4.4 ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ: ΣΥΝΔΥΑΣΜΟΣ ΓΛΩΣΣΙΚΩΝ ΕΥΡΗΜΑΤΩΝ ΚΑΙ ΔΕΙΚΤΩΝ ΠΟΛΙΤΙΚΗΣ 55
4.4.1 ΧΡΟΝΟΛΟΓΙΟ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ ΣΤΟΝ COVID-19 55
4.4.2 ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ 57
4.4.3 ΓΛΩΣΣΙΚΟΣ ΤΟΝΟΣ ΩΣ ΔΕΙΚΤΗΣ ΑΒΕΒΑΙΟΤΗΤΑΣ 59
4.4.4 ΣΥΜΠΕΡΑΣΜΑΤΑ ΣΥΓΚΡΙΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΝΟΜΙΣΜΑΤΙΚΗΣ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΗΣ ΠΟΛΙΤΙΚΗΣ 62
ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ 63
5.1 ΚΥΡΙΑ ΕΥΡΗΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ 63
5.2 ΣΥΜΒΟΛΗ ΕΡΓΑΣΙΑΣ 65
ΒΙΒΛΙΟΓΡΑΦΙΑ 67

ΕΥΡΕΤΗΡΙΟ ΔΙΑΓΡΑΜΜΑΤΩΝ-ΠΙΝΑΚΩΝ-ΣΧΗΜΑΤΩΝ:

Ε1. ΔΙΑΓΡΑΜΜΑΤΑ

ΔΙΑΓΡΑΜΜΑ 4.1.Γ: ΧΡΟΝΙΚΗ ΕΞΕΛΙΞΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ 2018-2024 (ΛΕΞΙΚΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ) 38

ΔΙΑΓΡΑΜΜΑ 4.1.Δ: ΘΕΜΑΤΙΚΕΣ ΚΑΤΗΓΟΡΙΕΣ ΑΝΑ ΕΤΟΣ (ΛΕΞΙΚΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ) 39

ΔΙΑΓΡΑΜΜΑ 4.1.ΣΤ: ΧΡΟΝΙΚΗ ΕΞΕΛΙΞΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (FINBERT).....41

ΔΙΑΓΡΑΜΜΑ 4.1.Η: ΣΥΓΚΡΙΣΗ FINBERT ΜΕ LEXICON (2018-2024).....43

ΔΙΑΓΡΑΜΜΑ 4.2.Α: ΚΥΡΙΑΡΧΕΣ ΛΕΞΕΙΣ ΑΝΑ TOPIC.....45

ΔΙΑΓΡΑΜΜΑ 4.2.Γ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2018.....47

ΔΙΑΓΡΑΜΜΑ 4.2.Δ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2019.....48

ΔΙΑΓΡΑΜΜΑ 4.2.Ε: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2020.....48

ΔΙΑΓΡΑΜΜΑ 4.2.ΣΤ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2021.....49

ΔΙΑΓΡΑΜΜΑ 4.2.Ζ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2022.....50

ΔΙΑΓΡΑΜΜΑ 4.2.Η: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2023.....51

ΔΙΑΓΡΑΜΜΑ 4.2.Θ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2024.....51

ΔΙΑΓΡΑΜΜΑ 4.2.Ι: ΠΟΣΟΣΤΟ ΕΜΦΑΝΙΣΗΣ ΠΛΗΘΩΡΙΣΜΟΥ ΣΤΑ ΘΕΜΑΤΑ ΑΝΑ ΧΡΟΝΙΑ 53

ΔΙΑΓΡΑΜΜΑ 4.4.Α: ΧΡΟΝΟΛΟΓΙΟ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ ΣΤΟΝ COVID-19(2020-2021) - ΝΟΜΙΣΜΑΤΙΚΕΣ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΕΣ ΠΑΡΕΜΒΑΣΕΙΣ 56

ΔΙΑΓΡΑΜΜΑ 4.4.Γ: ΣΥΣΧΕΤΙΣΗ ΜΕΣΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ FOMC ΚΑΙ ΕΒΔΟΜΑΔΙΑΙΩΝ ΑΙΤΗΣΕΩΝ ΑΝΕΡΓΙΑΣ (2020-2021) 60

Ε2. ΠΙΝΑΚΕΣ

ΠΙΝΑΚΑΣ 3.1.Α: ΧΡΟΝΙΚΗ ΔΙΑΡΘΡΩΣΗ ΠΕΡΙΟΔΩΝ ΑΝΑΛΥΣΗΣ(2018-2024) 21

ΠΙΝΑΚΑΣ 3.1.Β: ΠΕΡΙΓΡΑΦΙΚΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΟΛΟΥ ΔΕΔΟΜΕΝΩΝ.....23

ΠΙΝΑΚΑΣ 3.4.Α: ΚΡΙΤΗΡΙΑ ΓΙΑ ΑΡΙΘΜΟ TOPICS ΣΤΗΝ ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ 30

ΠΙΝΑΚΑΣ 3.4.Β: ΣΤΑΤΙΣΤΙΚΑ ΟΜΑΔΟΠΟΙΗΣΗΣ ΠΑΡΑΓΡΑΦΩΝ.....32

ΠΙΝΑΚΑΣ 3.4.Γ: ΚΡΙΤΗΡΙΑ ΓΙΑ ΑΡΙΘΜΟ TOPICS ΣΤΗΝ ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ33

ΠΙΝΑΚΑΣ 4.1.Α: ΚΑΤΑΝΟΜΗ ΛΕΞΕΩΝ ΑΝΑ ΚΑΤΗΓΟΡΙΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ) 36

ΠΙΝΑΚΑΣ 4.1.Β: ΕΤΗΣΙΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ(LEXICON).....37

ΠΙΝΑΚΑΣ 4.1.Ε: ΕΤΗΣΙΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ(FINBERT).....40

ΠΙΝΑΚΑΣ 4.2.Η: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ ΑΝΑ ΕΤΟΣ.....52

ΠΙΝΑΚΑΣ 4.1.Β: ΣΥΓΚΡΙΤΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΝΟΜΙΣΜΑΤΙΚΩΝ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΩΝ ΠΑΡΕΜΒΑΣΕΩΝ (2020-2021) 58

Ε3. ΣΧΗΜΑΤΑ

ΣΧΗΜΑ 4.1.Ζ: ΕΤΗΣΙΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ(FINBERT).....42

ΣΧΗΜΑ 4.2.Β: WORDCLOUD ΤΩΝ 5 ΘΕΜΑΤΩΝ ΤΗΣ ΑΝΑΛΥΣΗΣ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ45

ΛΕΞΙΚΟ ΟΡΩΝ:

Επιτροπή Ανοιχτής Αγοράς (Federal Open Market Committee - FOMC): Κύριο νομισματικό όργανο της Federal Reserve των ΗΠΑ, υπεύθυνο για τη διαμόρφωση της νομισματικής πολιτικής, κυρίως μέσω των επιχειρήσεων ανοιχτής αγοράς και του καθορισμού του στόχου για το επιτόκιο των federal funds. Αποτελείται από τα μέλη του Διοικητικού Συμβουλίου της Fed και τους προέδρους πέντε περιφερειακών τραπεζών της Fed.*(Bech & Lengwiler, 2022)*

Νομισματική Πολιτική (Monetary Policy):

Το σύνολο των ενεργειών μιας κεντρικής τράπεζας (όπως της Federal Reserve) που στοχεύουν στον έλεγχο της διαθεσιμότητας και του κόστους

του χρήματος και της πίστωσης, με απώτερο στόχο την προώθηση της σταθερότητας των τιμών, της πλήρους απασχόλησης και της βιώσιμης οικονομικής ανάπτυξης.*(Mishkin, 2022)*

Δείκτης Τιμών Προσωπικών Καταναλωτικών Δαπανών (Personal Consumption Expenditures Price Index - PCE):

Ο βασικός δείκτης τιμών που χρησιμοποιείται από τη FOMC για την επίσημη αξιολόγηση του επιπέδου πληθωρισμού. Μετρά τις μεταβολές στις τιμές των αγαθών και υπηρεσιών που πραγματικά καταναλώνουν τα νοικοκυριά.*(Federal Reserve, 2024)*

Συρρίκνωση (Contraction) / Ανάκαμψη (Expansion):

Φάσεις του οικονομικού κύκλου. Η συρρίκνωση χαρακτηρίζεται από πτώση της συνολικής οικονομικής δραστηριότητας, ενώ η ανάκαμψη από αύξηση.

Ποσοτική Χαλάρωση (Quantitative Easing - QE):

Μια μη συμβατική εργαλειοθήκη νομισματικής πολιτικής, κατά την οποία η κεντρική τράπεζα αγοράζει μακροπρόθεσμα οικονομικά στοιχεία (πχ. κυβερνητικά ομόλογα ή ομόλογα υψηλής ποιότητας) σε μεγάλη κλίμακα. Στόχος της είναι να μειώσει τα μακροπρόθεσμα επιτόκια, να αυξήσει τη ρευστότητα και να διεγείρει την οικονομική δραστηριότητα σε περιόδους που τα παραδοσιακά επιτόκια είναι κοντά στο μηδέν.(*Bernanke, 2020)*

Χρήματα Διακανονισμού (Federal Funds) / Επιτόκιο Διακανονισμού (Federal Funds Rate):

Τα χρήματα που οι εμπορικές τράπεζες δανείζουν η μία στην άλλη για βραχυπρόθεσμες ανάγκες ρευστότητας. Το επιτόκιο στο οποίο γίνονται αυτές οι συναλλαγές ονομάζεται επιτόκιο διακανονισμού (federal funds rate) και αποτελεί το βασικό εργαλείο πολιτικής της FOMC, καθώς επηρεάζει ένα ευρύ φάσμα άλλων επιτοκίων στην οικονομία.

Κύκλος Αυστηροποίησης (Tightening Cycle):

Μια περίοδος κατά την οποία μια κεντρική τράπεζα αυξάνει τα επιτόκια ή μειώνει τη ρευστότητα του χρηματοπιστωτικού συστήματος, με στόχο την κατάπνευση υπερθερμαινόμενων οικονομιών ή τον έλεγχο του πληθωρισμού.

Κείμενο ως Δεδομένο (Text as Data):

Μια μεθοδολογική προσέγγιση στις κοινωνικές επιστήμες που θεωρεί το γραπτό λόγο ως πηγή ποσοτικών πληροφοριών, μετατρέποντας μη δομημένα κείμενα σε δομημένες παρατηρήσεις που μπορούν να αναλυθούν με στατιστικές και οικονομετρικές μεθόδους.*(Gentzkow, Kelly, & Taddy, 2019)*

Ανάλυση Συναισθήματος (Sentiment Analysis) :

Υπολογιστική τεχνική που στοχεύει στην ταξινόμηση και ποσοτικοποίηση της υπόκρουσης (τόνου) ενός κειμένου (πχ. ως θετικό, αρνητικό, ουδέτερο). Στα οικονομικά κείμενα χρησιμοποιείται για την παρακολούθηση της οπτικής ή του επιπέδου αβεβαιότητας των θεσμικών φορέων.*(Araci, 2019)*

Λεξιλόγιο Συναισθήματος (Sentiment Dictionary/Lexicon):

Ένα προκαθορισμένο σύνολο λέξεων, κατηγοριοποιημένων ανά συναίσθημα (πχ. θετικό, αρνητικό, αβεβαιότητας). Στην εργασία αυτή χρησιμοποιείται το Loughran-McDonald Financial Sentiment Dictionary, ένα εξειδικευμένο λεξικό για οικονομικά κείμενα, το οποίο αποφεύγει τις προκαταλήψεις των γενικών λεξικών.*(Loughran & McDonald, 2011)*

Θεματική Ανάλυση (Topic Modeling):

Μια οικογένεια αλγορίθμων μη επιβλεπόμενης μάθησης για την αυτόματη ανάλυση μεγάλων συλλογών κειμένων, με στόχο την εύρεση κρυφών θεματικών προτύπων (topics). *(Blei, Ng, & Jordan, 2003)*

Λανθάνουσα Κατανομή Dirichlet (Latent Dirichlet Allocation - LDA): Το πιο διαδεδομένο στατιστικό μοντέλο γενεσιουργούς (generative model) για τη θεματική ανάλυση. Υποθέτει ότι κάθε έγγραφο είναι ένα μείγμα

διαφόρων θεμάτων και κάθε θέμα είναι μια κατανομή πιθανότητας σε ένα σύνολο λέξεων. *(Blei, Ng, & Jordan, 2003)*

BERT / FinBERT :

BERT (Bidirectional Encoder Representations from Transformers):

Μια αρχιτεκτονική νευρωνικού δικτύου μετασχηματιστή (transformer), προ-εκπαιδευμένη σε τεράστια κείμενα, ικανή να κατανοεί το συμφραζόμενο μιας λέξης αμφίδρομα. *(Devlin et al., 2018)*

FinBERT:

Μια προσαρμοσμένη έκδοση του BERT που έχει περάσει επιπλέον εκπαίδευση (fine-tuning) σε χρηματοοικονομικά κείμενα, βελτιστοποιώντας την απόδοσή της σε εργασίες όπως η ανάλυση συναισθήματος στον τομέα των οικονομικών. *(Araci, 2019)*

Προεπεξεργασία Κειμένου (Text Preprocessing):

Το σύνολο των τεχνικών καθαρισμού και τυποποίησης του ακατέργαστου κειμένου προς ανάλυση. Περιλαμβάνει: μετατροπή σε πεζά, αφαίρεση σημείων στίξης, διαχωρισμό σε λέξεις (tokenization), αφαίρεση κοινών λέξεων (stopwords) και αναγωγή σε λημματική μορφή (lemmatization).

Λέξεις χωρίς Ουσιαστική Σημασία (Stopwords):

Κοινές λέξεις (όπως "the", "is", "at", "which") που συνήθως αφαιρούνται κατά την προεπεξεργασία, καθώς εμφανίζονται τόσο συχνά που δεν προσφέρουν διακριτική πληροφορία για την ανάλυση. Στην παρούσα εργασία, το λεξικό stopwords επεκτάθηκε για να περιλαμβάνει και θεσμικούς/διαδικαστικούς όρους (πχ, "meeting", "participant").

Corpus:

Μια συστηματική συλλογή κειμένων, που χρησιμοποιείται ως το σύνολο δεδομένων για στατιστική ανάλυση. Στην εργασία αυτή, το corpus αποτελείται από τα 56 αρχεία FOMC Minutes για την περίοδο 2018-2024.

Δημοσιονομική Πολιτική (Fiscal Policy):

Οι ενέργειες της κυβέρνησης που αφορούν τη φορολογία και τις δαπάνες, με στόχο την επίτευξη μακροοικονομικών στόχων όπως η οικονομική σταθερότητα και η ανάπτυξη. Διαφέρει από τη νομισματική πολιτική που ασκείται από την κεντρική τράπεζα. *(Mishkin, 2022)*

Πρόγραμμα Προστασίας Μισθοδοσίας (Paycheck Protection Program - PPP):

Έκτακτο δημοσιονομικό μέτρο που εφαρμόστηκε στις ΗΠΑ κατά την πανδημία COVID-19, με στόχο την παροχή δανείων συγχώρησης σε μικρές επιχειρήσεις για να καλύψουν έξοδα όπως οι μισθοί και το ενοίκιο. *(U.S. Department of the Treasury, 2020)*

Συσχέτιση (Correlation):

Στατιστικός όρος που μετρά την ισχύ και την κατεύθυνση της γραμμικής σχέσης μεταξύ δύο μεταβλητών. Η τιμή του (r) κυμαίνεται από -1 (τέλεια αρνητική συσχέτιση) έως +1 (τέλεια θετική συσχέτιση). Μια τιμή κοντά στο 0 υποδεικνύει ασθενή ή ανύπαρκτη γραμμική σχέση.*(Field, 2018)*

Χρονολόγιο Πολιτικής (Policy Timeline):

Μια γραφική ή αφηγηματική αναπαράσταση της χρονικής ακολουθίας σημαντικών πολιτικών αποφάσεων και γεγονότων, που επιτρέπει την ανάλυση της προτεραιότητας, της συγχρονισμένης δράσης και της εξέλιξης των στρατηγικών.

Δείκτης Επίδρασης Πολιτικής (Policy Impact Metric):

Ένας ποσοτικός ή ποιοτικός δείκτης που επιχειρεί να μετρήσει ή να κατατάξει την αποτελεσματικότητα ή την έκταση μιας πολιτικής παρέμβασης. Στην ποσοτική ανάλυση μπορεί να βασίζεται σε αλλαγές σε οικονομικούς δείκτες, ενώ σε ποιοτική ανάλυση μπορεί να βασίζεται σε εμπεριστατωμένη εκτίμηση. *(Gentzkow et al., 2019)*

ΚΕΦΑΛΑΙΟ 1: ΕΙΣΑΓΩΓΗ

Οι κεντρικές τράπεζες, και ειδικότερα η Federal Reserve των Ηνωμένων Πολιτείων, διαδραματίζουν καθοριστικό ρόλο στη διαχείριση της μακροοικονομικής σταθερότητας. Η επικοινωνία τους ιδίως μέσω των επίσημων εγγραφών, όπως των Πρακτικών (Minutes) της Επιτροπής Ανοιχτής Αγοράς (Federal Open Market Committee, FOMC), αποτελεί κανάλι μετάδοσης πληροφοριών προς τις αγορές, τους οικονομικούς φορείς και το ευρύ κοινό *(Hansen et al., 2018)*. Η ερμηνεία αυτών των κειμένων βασίζεται παραδοσιακά σε ποιοτική ανάλυση, ωστόσο οι πρόσφατες εξελίξεις στην Επεξεργασία Φυσικής Γλώσσας (Natural Language Processing – NLP) έχουν επαναπροσδιορίσει τις δυνατότητες της ποσοτικής ανάλυσης του θεσμικού λόγου με την προσέγγιση «Κείμενο ως Δεδομένο» (Text as Data) *(Gentzkow et al., 2019)*.

Η περίοδος 2018-2024, η οποία θα μελετηθεί στη παρούσα εργασία, χαρακτηρίζεται από μια σειρά έντονων και διαδοχικών διαταραχών: από τη σχετική ηρεμία της προ-πανδημικής εποχής στο σεισμικό γεγονός της πανδημίας COVID-19, την ταχεία πληθωριστική έκρηξη που ακολούθησε, τις τραπεζικές αναταραχές του 2023, και τις πρώτες ενδείξεις σταθεροποίησης το 2024. Μέσα σε αυτό το ταραγμένο πλαίσιο, η FOMC αντιμετώπισε την ανάγκη για ταχεία και πολύπλοκη προσαρμογή της πολιτικής της. Τα πρακτικά των συνεδριάσεών της αποτελούν το πλουσιότερο δημόσιο αρχείο των εσωτερικών συζητήσεων, της ανάλυσης των κινδύνων και της λογικής λήψης αποφάσεων κατά τη διάρκεια αυτής της κρίσιμης περιόδου *(Shin, 2021)*.

Παρά το αυξανόμενο ακαδημαϊκό ενδιαφέρον για την ανάλυση του λόγου της FOMC, παρατηρείται ένα σημαντικό ερευνητικό κενό όσον αφορά την πολυμεθοδική και διαχρονική εξέταση τόσο του περιεχομένου (τί) όσο και του συναισθηματικού τόνου (πώς) των συζητήσεών της. Προηγούμενες σημαντικές έρευνες εστιάζουν είτε στη θεματική ανάλυση *(πχ. Shin, 2021)*, είτε στη μέτρηση του συναισθήματος *(πχ. Hansen et al., 2018)*. Ωστόσο, ο συστηματικός συνδυασμός κλασικών λεξικολογικών προσεγγίσεων *(Loughran & McDonald, 2011)* με προηγμένα γλωσσικά μοντέλα *(FinBERT, Araci, 2019)* στον ίδιο πυρήνα δεδομένων παραμένει σπάνιος και παράλληλα ακολουθεί την εξέλιξη αυτών των διαστάσεων σε ένα εκτεταμένο χρονικό ορίζοντα που καλύπτει πολλαπλές φάσεις του οικονομικού κύκλου.

Στόχος της παρούσας διπλωματικής εργασίας είναι να συμβάλει γεμίζοντας αυτό το ερευνητικό κενό. Μέσω μιας πολυμεθοδικής

ανάλυσης των FOMC Minutes για την περίοδο Ιανουαρίου 2018 έως Δεκεμβρίου 2024, επιχειρείται:

1. Η μέτρηση και η ανάλυση του συναισθηματικού τόνου των συζητήσεων, χρησιμοποιώντας συγκριτικά δύο διαφορετικές προσεγγίσεις Ανάλυσης Συναισθήματος (Sentiment Analysis): το εξειδικευμένο χρηματοοικονομικό λεξικό Loughran-McDonald και το προ-εκπαιδευμένο γλωσσικό μοντέλο FinBERT.
2. Ο εντοπισμός και η ποσοτικοποίηση των θεματικών προτεραιοτήτων της Επιτροπής και η χρονική τους εξέλιξη, εφαρμόζοντας Θεματική Ανάλυση (Topic Modeling) με το μοντέλο Latent Dirichlet Allocation (LDA) σε δύο επίπεδα: ανά παράγραφο και ανά έτος.
3. Η ερμηνεία των μεταβολών τόσο στα θέματα όσο και στο συναίσθημα σε σχέση με το ευρύτερο ιστορικό και οικονομικό πλαίσιο κάθε υποπεριόδου (προ-πανδημική, πανδημία, πληθωριστική κρίση, σταθεροποίηση).
4. Η διερεύνηση της σχέσης και του συντονισμού μεταξύ της γλωσσικής έκφρασης των πρακτικών (ευρήματα των στόχων 1-3) και των πραγματικών νομισματικών και δημοσιονομικών πολιτικών μέτρων που ελήφθησαν κατά τη διάρκεια της πανδημίας COVID-19, αξιολογώντας έτσι την πολιτική ανταπόκριση.

Η εργασία αυτή συνεισφέρει τόσο στη βελτίωση της εμπειρικής κατανόησης της νομισματικής πολιτικής και της επικοινωνίας της Fed σε καιρούς κρίσης, όσο και στη μεθοδολογική συζήτηση σχετικά με τις βέλτιστες πρακτικές ανάλυσης πολύπλοκων θεσμικών κειμένων. Τα ευρήματα αναμένεται να καταδείξουν πώς η γλώσσα των πρακτικών λειτουργεί ως ευαίσθητος δείκτης οικονομικής αβεβαιότητας και πολιτικής προσαρμογής.

Η δομή της εργασίας είναι η εξής: Το Κεφάλαιο 2 παρουσιάζει το θεωρητικό και βιβλιογραφικό πλαίσιο. Το Κεφάλαιο 3 εξηγεί λεπτομερώς τη μεθοδολογία συλλογής και ανάλυσης των δεδομένων. Το Κεφάλαιο 4 παρουσιάζει και ερμηνεύει τα βασικά αποτελέσματα της ανάλυσης. Το Κεφάλαιο 5 συνοψίζει τα βασικά συμπεράσματα και τις πιθανές επιπτώσεις.

ΚΕΦΑΛΑΙΟ 2: ΘΕΩΡΗΤΙΚΟ ΚΑΙ ΒΙΒΛΙΟΓΡΑΦΙΚΟ ΠΛΑΙΣΙΟ

Η παρούσα διπλωματική εντάσσεται στο ευρύτερο και ταχέως αναπτυσσόμενο πεδίο έρευνας της υπολογιστικής γλωσσολογίας και της ανάλυσης κειμένου με την προσέγγιση «Κείμενο ως δεδομένο» (text as data) στις οικονομικές επιστήμες. Στόχος αυτού του Κεφαλαίου είναι να παρουσιάσει το θεωρητικό και βιβλιογραφικό πλαίσιο πάνω στο οποίο στηρίζεται η εμπειρική ανάλυση. Το Κεφάλαιο διέπεται από τρεις βασικούς άξονες:

A. Τη μεθοδολογική επανάσταση της προσέγγισης «κείμενο ως δεδομένο»,
B. Τις τεχνικές Ανάλυσης Συναισθήματος(Sentiment Analysis)με έμφαση στην εξέλιξη από τα στατικά λεξικά στα προηγμένα γλωσσικά μοντέλα και
C. Τη Θεματική Ανάλυση(topic modeling) μέσω του μοντέλου Latent Dirichlet Allocation(LDA).

Αρχικά, θα αναλυθεί το θεμελιώδες έργο του *Gentzkow, Kelly και Taddy (2019)*, το οποίο συστηματοποιεί τις δυνατότητες και τις προκλήσεις της χρήσης κειμένου ως ποσοτικών δεδομένων στην οικονομική επιστήμη. Στη συνέχεια, η βιβλιογραφική ανασκόπηση θα επικεντρωθεί στις δύο βασικές προσεγγίσεις για την Ανάλυση Συναισθήματος. Ξεκινώντας με το σημασιολογικό λεξικό Loughran και McDonald (2011), το οποίο καθιέρωσε την αναγκαιότητα χρήσης εξειδικευμένων λεξικών για την ακριβή ανάλυση οικονομικών κειμένων, αποφεύγοντας τις σημαντικές προκαταλήψεις των γενικών λεξικών. Συνεχίζοντας, με τη πρόσφατη εξέλιξη προς μοντέλα βασισμένα σε εξελιγμένες υπολογιστικές τεχνικές γλωσσικής ανάλυσης. Κύριο εργαλείο αυτής της προσέγγισης είναι το FinBERT *(Araci, 2019)*, ένα προ-εκπαιδευμένο γλωσσικό μοντέλο ικανό να κατανοεί τα συμφραζόμενα και τις σύνθετες "σημασιολογικές αποχρώσεις" στα χρηματοοικονομικά κείμενα.

Ο τρίτος άξονας αφορά την αυτόματη εξαγωγή θεμάτων (topic modeling). Θα παρουσιαστεί το θεωρητικό υπόβαθρο του μοντέλου Latent Drichlet Allocation(LDA) όπως διατυπώθηκε από τους *Blei, Ng και Jordan (2003)*,

οι οποίοι θέσπισαν το θεωρητικό υπόβαθρο για την ανάλυση μεγάλων συλλογών κειμένων. Τέλος, θα αναλυθεί το εφαρμοσμένο έργο του *Shin(2021)*, το οποίο εφαρμόζει τη μέθοδο LDA στα Πρακτικά της FOMC για να καταγράψει τη μετατόπιση των συζητήσεων γύρω από την πανδημία COVID-19, αποτελώντας σχετικό πρότυπο για την παρούσα έρευνα.

Μετά από αυτή την ανασκόπηση θα διαπιστωθεί ότι ενώ υπάρχουν σημαντικές προηγούμενες μελέτες που εφαρμόζουν μεμονωμένες τεχνικές, υπάρχει ένα ερευνητικό κενό όσον αφορά τον συνδυασμό της θεματικής ανάλυσης με την πολυμεθοδική ανάλυση συναισθήματος για τη μελέτη της επικοινωνίας της FOMC. Η παρούσα εργασία στοχεύει να συμβάλει γεμίζοντας αυτό το κενό, προσφέροντας μια πολυδιάστατη κατανόηση τόσο των θεματικών προτεραιοτήτων όσο και του συναισθηματικού τόνου στις αποφάσεις της Επιτροπής, σε μια περίοδο έντονων οικονομικών σοκ (2018-2024).

2.1 ΠΡΟΣΕΓΓΙΣΗ «ΚΕΙΜΕΝΟ ΩΣ ΔΕΔΟΜΕΝΟ»

Η ραγδαία ανάπτυξη της υπολογιστικής ισχύος και των τεχνικών επεξεργασίας φυσικής γλώσσας (NLP) έχει επιφέρει μια θεμελιώδη μετατόπιση στον τρόπο με τον οποίο οι οικονομικές επιστήμες προσεγγίζουν πληροφορίες που θεωρούνται προσιτές μόνο μέσω ποιοτικής ερμηνείας. Η προσέγγιση «Text as Data» (Κείμενο ως δεδομένο) ενσωματώνει τον γραπτό λόγο στην ποσοτική ανάλυση, μετατρέποντας μη-δομημένα κείμενα (όπως ομιλίες, αναφορές και πρακτικά συνεδριάσεων) σε δομημένες παρατηρήσεις για στατιστική επεξεργασία. Αυτή η μεθοδολογική "επανάσταση" επιτρέπει την εξαγωγή πληροφορίων από τεράστιες συλλογές δεδομένων, σε κλίμακα και αντικειμενικότητα που προηγουμένως ήταν αδιανόητη.

Το θεμελιώδες έργο των *Gentzkow, Kelly και Taddy (2019)* με το τίτλο «Text as Data» παρέχει την πιο ολοκληρωμένη επισκόπηση αυτού του πεδίου. Οι συγγραφείς ορίζουν τρεις βασικές προκλήσεις της ποσοτικής ανάλυσης κειμένου:

1. **της Αντιπροσώπευσης (Representation):** Η μετατροπή λέξεων σε αριθμητική αναπαράσταση για στατιστικά/οικονομετρικά μοντέλα.
2. **Της Εξαγωγής (Discovery):** Η μετατροπή λέξεων σε αριθμητική αναπαράσταση για στατιστικά/οικονομετρικά μοντέλα
3. **Της Εκτίμησης (Measurement):** Η ποσοτικοποίηση χαρακτηριστικών όπως το συναίσθημα ή η αβεβαιότητα.

Όπως τονίζουν οι συγγραφείς, "η επιλογή μεθόδου εξαρτάται βαθιά από το ερευνητικό ερώτημα". *(Gentzkow et al., 2019, p. 5)*. Για την παρούσα έρευνα, οι κύριες προκλήσεις που καλούνται να ανταπεξέλθουν είναι η Εξαγωγή, στην οποία κυριαρχεί η Θεματική Ανάλυση με τεχνικές όπως το LDA και η Εκτίμηση στην όποια κυριαρχεί η Ανάλυση Συναισθήματος με τεχνικές από λεξικά έως νευρωνικά δίκτυα.

Η κεντρική συμβολή αυτού του πλαισίου είναι η έμφαση που δίνει στην προσαρμογή της μεθόδου στο ερευνητικό ερώτημα και στο ειδικό λεξιλόγιο του υπό μελέτη πεδίου. Αυτή η αρχή δικαιολογεί την επιλογή εξειδικευμένων εργαλείων για την ανάλυση των Πρακτικών μιας Κεντρικής Τράπεζας και αποτελεί την αφετηρία για τη διερεύνηση των δύο βασικών μεθοδολογιών αυτής της εργασίας: της Ανάλυσης Συναισθήματος και της Θεματικής Ανάλυσης. Μέσω αυτών, η προσέγγιση «Κείμενο ως Δεδομένο» επιτρέπει μια ολιστική κατανόηση τόσο του τρόπου (πώς) όσο και του περιεχομένου (τί) της επικοινωνίας της FOMC.

2.2 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (SENTIMENT ANALYSIS)

Η Ανάλυση Συναισθήματος αποσκοπεί στην ποσοτικοποίηση του υποκειμενικού τόνου ή της συναισθηματικής χροιάς ενός κειμένου. Στα οικονομικά, η εφαρμογή της έχει αποκαλύψει πολύτιμες πληροφορίες για την επικοινωνιακή στρατηγική εταιριών, την αντίδραση των αγορών και τις προοπτικές των θεσμικών φορέων. Η εξέλιξη των μεθόδων ακολούθησε δύο βασικούς άξονες: Τις στατικές, λεξικολογικές προσεγγίσεις και τις δυναμικές προσεγγίσεις βασισμένες σε νευρωνικά δίκτυα και κατανόηση του συμφραζομένου.

2.2.1 ΛΕΞΙΚΟΛΟΓΙΚΕΣ ΠΡΟΣΕΓΓΙΣΕΙΣ

Οι πρώτες και πιο διαδεδομένες μέθοδοι ανάλυσης συναισθήματος βασίζονται σε προ-καθαρισμένα λεξικά (dictionaries). Αυτά τα λεξικά αντιστοιχίζουν λέξεις σε κατηγορίες συναισθήματος (πχ. Θετικές, αρνητικές, αβεβαιότητας). Η τεχνική που ακολουθείται είναι απλή: Η συχνότητα εμφάνισης λέξεων κάθε κατηγορίας σε ένα κείμενο υπολογίζεται και χρησιμοποιείται για την παραγωγή ενός συνόλου βαθμολόγησης.

Ωστόσο, η εφαρμογή γενικών λεξικών συναισθήματος, όπως το ευρέως χρησιμοποιούμενο Harvard General Inquirer (Harvard-IV-4) αποδείχθηκε προβληματική. Η μελέτη που αποτύπωσε αυτό το θεμελιώδες λάθος και επαναπροσδιόρισε τις βέλτιστες πρακτικές είναι το έργο των *Loughran και McDonald (2011)* με τίτλο *«When is a Liability Not a Liability? Textual Analysis, Dictionaries, and 10-Ks»*.

Το κεντρικό επιχείρημα του paper είναι ότι οι λέξεις αποκτούν διαφορετικό νόημα ανάλογα με το πεδίο χρήσης τους. Οι συγγραφείς επισήμαναν ότι σε ετήσιες οικονομικές εκθέσεις εταιριών (10-Ks), "Περίπου το 75% των λέξεων που τα γενικά λεξικά χαρακτηρίζει ως «αρνητικές» είναι στην πραγματικότητα ουδέτεροι ή θετικοί τεχνικοί όροι στον οικονομικό λόγο" *(Loughran & McDonald, 2011, p.35)*.

Για να αντιμετωπιστεί αυτό το σφάλμα, οι Loughran και McDonald (2011) δημιούργησαν ένα εξειδικευμένο χρηματοοικονομικό λεξικό μέσω επαγωγικής ανάλυσης ενός μεγάλου corpus οικονομικών αναφορών. Αυτό το λεξικό, ταξινομώντας τις λέξεις βάσει της πραγματικής τους χρήσης στον οικονομικό λόγο, αποφεύγει τις προκαταλήψεις των γενικών λεξικών και έχει αποδειχθεί ότι διαθέτει σημαντικά μεγαλύτερη προβλεπτική ισχύ για πραγματικά οικονομικά αποτελέσματα. Η χρήση αυτού του λεξικού αποτελεί, επομένως μια κρίσιμη απόφαση μεθοδολογικής ακρίβειας που αποφεύγει τις προκαταλήψεις των γενικών λεξικών.

2.2.2 ΠΡΟΣΕΓΓΙΣΕΙΣ ΒΑΣΙΣΜΕΝΕΣ ΣΕ ΝΕΥΡΩΝΙΚΑ ΔΙΚΤΥΑ - ΕΠΑΝΑΣΤΑΣΗ ΤΟΥ FINBERT

Παρά τη σημασία τους οι λεξικολογικές προσεγγίσεις, ακόμη και οι εξειδικευμένες παρουσιάζουν περιορισμούς στην κατανόηση του συμφραζόμενο, όπως τη διπλή άρνηση (πχ. "Η ανάπτυξη *δεν ήταν* αδύναμη"), τη μεταφορική γλώσσα και την σημασία ολόκληρων προτάσεων ή παραγράφων. Για την αντιμετώπιση αυτών των περιορισμών, η έρευνα στράφηκε σε προ-εκπαιδευμένα γλωσσικά μοντέλα βασισμένα σε αρχιτεκτονικές Μετασχηματιστών (Transformers).

Η επανάσταση ξεκίνησε με το μοντέλο BERT (Bidirectional Encoder Representations from Transformers) *(Devlin et al., 2018)*, το οποίο είναι εκπαιδευμένο σε τεράστια κείμενα ,ώστε να μάθει πως η σημασία μιας λέξης αλλάζει με βάση τις λέξεις που την περιβάλλουν σε αμφίδρομη κατεύθυνση. Ωστόσο, δεν είναι βέλτιστα προσαρμοσμένο για τον ιδιαίτερο λόγο των χρηματοοικονομικών.

Την ανάγκη για περαιτέρω εξειδίκευση την καλύπτει το μοντέλο FinBERT *Araci(2019)* . Στο paper *«FinBERT: Financial Sentiment Analysis with Pre-trained Language Models», ο Araci* παρουσιάζει ένα BERT μοντέλο που έχει περάσει επιπλέον εκπαίδευση(fine-tuning) σε ένα τεράστιο corpus χρηματοοικονομικών κειμένων. Το βασικό του πλεονέκτημα είναι η ικανότητα του να αναλύει ολόκληρες προτάσεις και παραγράφους, εκτιμώντας πιθανότητες για κάθε κατηγορία συναισθήματος (θετικό, αρνητικό, ουδέτερο). Αυτό του επιτρέπει να κατανοεί σύνθετες σημασιολογικές αποχρώσεις, να διακρίνει αναμεμειγμένα συναισθήματα και να αναγνωρίζει ότι, για παράδειγμα, η φράση *"Η αύξηση των τιμών ενέχει κινδύνους"* φέρει αρνητικό τόνο, παρά τη λέξη *"αύξηση"*.

Στην παρούσα έρευνα, η χρήση του FinBERT συμπληρώνει τη λεξικολογική προσέγγιση (Loughran-McDonald). Αυτός ο συνδυασμός επιτρέπει, από τη μία τη διαφάνεια και επαναληψιμότητα της στατικής λεξικής ανάλυσης και από την άλλη το βάθος και την ευελιξία της συμφραζομενικής ερμηνείας. Η σύγκριση των αποτελεσμάτων και των δύο μεθόδων στο Κεφάλαιο 4 θα προσφέρει μια ολιστική και ισχυρή κατανόηση του συναισθηματικού τόνου στα FOMC Minutes.

2.3 ΘΕΜΑΤΙΚΗ ΑΝΑΛΥΣΗ (TOPIC MODELING) ΚΑΙ ΜΟΝΤΕΛΟ LATENT DIRICHLET ALLOCATION (LDA)

Η ανάλυση του συναισθήματος μετράει τον «τρόπο» (το πώς) της επικοινωνίας, η Θεματική Ανάλυση (Topic Modeling) στοχεύει στο να αποκαλύψει το «περιεχόμενο» (το τί). Σε μεγάλες συλλογές κειμένων όπως τα πρακτικά πολλών ετών της FOMC, δεν είναι εφικτή η ανάγνωση όλων και η εξαγωγή των κύριων θεματικών αξόνων που απασχολούν τους συμμετέχοντες. Η Θεματική Ανάλυση παρέχει ένα σύνολο υπολογιστικών μεθόδων για την αυτόματη ομαδοποίηση και εξαγωγή των επικρατούντων θεμάτων(topics) που διατρέχουν τα κείμενα, χωρίς προκαθορισμένη ετικέτα ή επίβλεψη.

Η βασική μέθοδος για αυτό το σκοπό είναι το Latent Dirichlet Allocation (LDA) (*Blei, Ng & Jordan,2003). Το* LDA βασίζεται στην απλή αλλά ισχυρή υπόθεση ότι κάθε έγγραφο είναι ένα μείγμα διαφόρων θεμάτων και κάθε θέμα είναι μια κατανομή πιθανοτήτων σε λέξεις. Μέσω αυτού του πλαισίου, το LDA είναι ο αλγόριθμος που εντοπίζει τα άυλα θέματα που διατρέχουν το corpus και ποσοτικοποιεί τη συμβολή τους σε κάθε κείμενο.

Στην παρούσα έρευνα το LDA αποτελεί το βασικό εργαλείο για την απάντηση σε ένα από τα ερευνητικά ερωτήματα σχετικά με τις θεματικές προτεραιότητες της FOMC. Η εφαρμογή του ακολουθεί πιστά το θεωρητικό πλαίσιο των *Blei et al.(2003).* Συγκεκριμένα, υλοποιείται σε δύο επίπεδα:

- Στο επίπεδο της παραγράφου: Για τον εντοπισμό των σταθερών θεματικών πυλώνων που χαρακτηρίζουν τις συζητήσεις της FOMC σε όλη την υπό μελέτη περίοδο (2018-2024), ανεξάρτητα από τα ειδικά γεγονότα κάθε έτους.
- Στο επίπεδο ετών: Για την ανίχνευση της χρονικής εξέλιξης και μετατόπισης των προτεραιοτήτων της Επιτροπής, επιτρέποντας την παρατήρηση της άμεσης ανταπόκρισής της σε σημαντικά σοκ, όπως η πανδημία COVID-19.

Η επιλογή του αριθμού των θεμάτων(k) σε κάθε ανάλυση βασίζεται σε προσαρμοστική στρατηγική που εξαρτάται από τον όγκο των δεδομένων, με στόχο την παραγωγή ερμηνεύσιμων και στατιστικά σταθερών αποτελεσμάτων. Αυτή η διπλή προσέγγιση επιτρέπει μια ολιστική κατανόηση τόσο των βασικών όσο και των μεταβαλλόμενων στοιχείων της επικοινωνίας της FOMC.

2.4 ΠΡΟΗΓΟΥΜΕΝΕΣ ΕΦΑΡΜΟΓΕΣ ΣΕ FOMC MINUTES

Η ανάλυση της επικοινωνίας της FOMC είναι ένα ενεργό πεδίο μελέτης, καθώς τα πρακτικά αποτελούν τον πλουσιότερο και λεπτομερέστερο δημόσιο αρχείο των εσωτερικών συζητήσεων και της διαδικασίας λήψης αποφάσεων της Κεντρική Τράπεζας των ΗΠΑ.

Ένα άμεσα συναφές και μεθοδολογικά καθοδηγητικό παράδειγμα είναι η μελέτη του *Shin (2021) με τίτλο «Topic Modeling of FOMC Minutes: Evolution of Discussions Before and After the COVID-19 Outbreak».* Η οποία εφάρμοσε θεματική ανάλυση (LDA) στα FOMC Minutes για να χαρτογραφήσει τη μετατόπιση των συζητήσεων της Επιτροπής πριν και μετά την έναρξη της πανδημίας COVID-19.

Τα βασικά ευρήματα του *Shin (2021*) που διαμορφώνουν το πλαίσιο της παρούσας έρευνας είναι τρία:

1. **Επιβεβαίωση της Εφαρμοσιμότητας:** Αποδεικνύει ότι το LDA είναι ένα ισχυρό και ερμηνεύσιμο εργαλείο για την εξαγωγή θεμάτων από τα πρακτικά.

2. **Τεκμηρίωση Θεματικής Μετατόπισης:** Καταγράφει μια σαφή μεταβολή στις προτεραιότητες με την έλευση της πανδημίας, όπου έντονα νέα θέματα (όπως η χρηματοπιστωτική σταθερότητα) αντικατέστησαν σε προτεραιότητα παλαιότερα.

3. **Δημιουργία ενός Πρωτότυπου:** Η χρονική ανάλυση θεμάτων (temporal topic analysis) του Shin λειτουργεί ως άμεσα συναφές πρότυπο για την παρούσα έρευνα, υποδεικνύοντας ότι αυτή η μέθοδος είναι ουσιαστική για την κατανόηση του πώς η FOMC ανταποκρίνεται σε εξωγενείς σοκ.

Έχοντας παρουσιαστεί αυτό το κρίσιμο υπόδειγμα, η βιβλιογραφική ανασκόπηση προχωρά στο συνολικό συμπέρασμα και στον εντοπισμό του ερευνητικού κενού που καλείται να καλύψει η παρούσα μελέτη.

2.5 ΣΥΜΠΕΡΑΣΜΑΤΑ ΒΙΒΛΙΟΓΡΑΦΙΚΗΣ ΑΝΑΣΚΟΠΗΣΗΣ

Η βιβλιογραφική ανασκόπηση που προηγήθηκε αποτυπώνει την ταχεία εξέλιξη των μεθόδων ανάλυσης κειμένου και την εφαρμογή τους στη μελέτη της νομισματικής επικοινωνίας των κεντρικών τραπεζών. Μπορούν να εξαχθούν τα εξής βασικά συμπεράσματα:

- Η προσέγγιση «Κείμενο ως Δεδομένο» έχει θεμελιώσει ένα νέο παράδειγμα έρευνας, επιτρέποντας την ποσοτικοποίηση και συστηματική ανάλυση τεράστιων όγκων κειμένου.
- Το πεδίο της ανάλυσης συναισθήματος, έχει εξελιχθεί από τα εξειδικευμένα οικονομικά λεξικά (*Loughran & McDonald, 2011)* στα προηγμένα γλωσσικά μοντέλα *(FinBERT, Araci, 2019)*, που προσφέρουν ανώτερη κατανόηση συμφραζομένου και σύνθετων σημασιολογικών αποχρώσεων.
- Η Θεματική Ανάλυση μέσω LDA παραμένει το βασικό μοντέλο για την αυτόματη εξαγωγή θεμάτων και έχει αποδειχθεί αποτελεσματικά εφαρμόσιμη στην ανάλυση των FOMC Minutes, όπως επιδεικνύει το έργο του *Shin (2021).*

Ωστόσο, μια κριτική σύνθεση αυτής της βιβλιογραφίας αποκαλύπτει ένα σαφές ερευνητικό κενό, το οποίο καθορίζει τη συμβολή της παρούσας διπλωματικής εργασίας. Αυτό το κενό έχει να κάνει ως προς την πολυμεθοδικότητα. Παράλληλα, υπάρχει σχετική απουσία ερευνών που να χρησιμοποιούν συστηματικά τα γλωσσικά ευρήματα (Sentiment, topics) ως

γέφυρα για την ποσοτική αξιολόγηση της ίδιας της διαδικασίας λήψης πολιτικών αποφάσεων και του συντονισμού της με άλλους πολιτικούς τομείς, όπως τη δημοσιονομική πολιτική. Ειδικά σε περιόδους έκτακτων κρίσεων όπως η πανδημία COVID-19, η σύνδεση ανάμεσα στη θεσμική ρητορική και τις πραγματικές πολιτικές ενέργειες παραμένει ένα πεδίο που αξίζει περαιτέρω διερεύνησης. Σημαντικές έρευνες όπως του *Shin (2021)*, επικεντρώνονται αποκλειστικά ή κυρίως στη θεματική ανάλυση (το «τι»). Αυτό αφήνει ακάλυπτη τη διερεύνηση του συναισθηματικού τόνου (το «πώς») που συνοδεύει αυτές τις συζητήσεις. Η πληρέστερη κατανόηση της επικοινωνίας της FOMC απαιτεί τη συνύπαρξη και των δύο διαστάσεων.

Η παρούσα διπλωματική εργασία προτείνει τον συνδυασμό της Ανάλυσης Θεμάτων (LDA) σε δύο επίπεδα (παράγραφος και έτος), με μια διπλή και συγκριτική Ανάλυση Συναισθήματος (Λεξιλογική Loughran-McDonald και Νευρωνικό Δίκτυο FinBERT), η παρούσα μελέτη στοχεύει να προσφέρει μια πολυδιάστατη και χρονικά ευαίσθητη κατανόηση της επικοινωνίας της FOMC για την περίοδο 2018-2024. Αυτή η προσέγγιση επιτρέπει να απαντηθούν ταυτόχρονα τρία ερωτήματα:

1. Ποια ήταν τα βασικά και μεταβαλλόμενα θέματα συζήτησης;

2. Ποιος ήταν ο συναισθηματικός τόνος αυτών των συζητήσεων και πώς διακυμαίνεται;

3. Πώς συμβάλλουν και συμπληρώνονται οι διαφορετικές μέθοδοι ανάλυσης συναισθήματος στην ερμηνεία αυτών των δεδομένων;

Με αυτόν τον τρόπο, η εργασία συμβάλλει τόσο στη βελτίωση της εμπειρικής κατανόησης της νομισματικής πολιτικής, όσο και στη μεθοδολογική συζήτηση σχετικά με την ανάλυση θεσμικών κειμένων.

ΚΕΦΑΛΑΙΟ 3: ΜΕΘΟΔΟΛΟΓΙΑ

3.1 ΣΥΛΛΟΓΗ ΔΕΔΟΜΕΝΩΝ

Για την παρούσα διπλωματική εργασία, συλλέχθηκαν και αναλύθηκαν τα πρακτικά των συνεδριάσεων της Federal Open Market Committee (FOMC Minutes). Σε αυτό το κεφάλαιο θα αναλυθεί η μεθοδολογία της ανάλυσης και η εφαρμογή της με την χρήση της PYTHON. Η ανάλυση επικεντρώθηκε στην περίοδο από τον Ιανουάριο του 2018 έως τον Δεκέμβριο του 2024. Η συγκεκριμένη χρονική περίοδος επιλέχθηκε με στόχο την ανάλυση της πανδημίας COVID-19. Χωρίστηκε σε τέσσερις φάσεις ώστε να αναλυθεί και η περίοδος πριν αλλά και μετά την πανδημία . Με στόχο τον εντοπισμό κοινών σημείων αλλά και διαφορών στον λόγο των ανακοινώσεων της Federal. Οι τέσσερις φάσεις που ορίστηκαν για την ανάλυση, με βάση τις κυρίαρχες μακροοικονομικές συνθήκες κάθε υποπεριόδου, συνοψίζονται στον Πίνακα 3.1.Α και περιγράφονται αναλυτικά στη συνέχεια.

ΠΙΝΑΚΑΣ 3.1.Α: ΧΡΟΝΙΚΗ ΔΙΑΡΘΡΩΣΗ ΠΕΡΙΟΔΩΝ ΑΝΑΛΥΣΗΣ(2018-2024)

ΕΤΗ	ΟΝΟΜΑΣΙΑ ΠΕΡΙΟΔΟΥ
2018-2019	Προ-πανδημική Περίοδος
2020-2021	Περίοδος Πανδημίας COVID-19
2022-2023	Μετά-Πανδημική και Πληθωριστική Περίοδος
2024	Περίοδος σταθεροποίησης

1. **Προ-πανδημική περίοδος (2018-2019):** Αφορά την περίοδο πριν την εμφάνιση της πανδημίας Covid-19. Αυτή η περίοδος χαρακτηρίζεται από σχετική οικονομική σταθερότητα, καθώς και χαμηλά ποσοστά ανεργίας και πληθωρισμού. Μελετήθηκε με σκοπό να παρατηρηθεί αν υπήρξε κάποια απότομη αλλαγή με την εμφάνιση του COVID-19.

2. **Περίοδος πανδημίας COVID-19(2020-2021):** Περιλαμβάνει την απότομη μείωση της οικονομικής δραστηριότητας που εμφανίστηκε το πρώτο εξάμηνο του 2020, τις έκτακτες νομισματικές και δημοσιονομικές παρεμβάσεις που χρειάστηκε να ληφθούν, καθώς και την αρχική ανάκαμψη του 2021 που οφειλόταν στην άφιξη των εμβολίων.

3. **Μετά-πανδημική και Πληθωριστική περίοδος (2022-2023):** Αφού ξεπεράστηκε η πανδημία, η οικονομία αντιμετώπισε ραγδαία άνοδο του πληθωρισμού, με αποτέλεσμα την έναρξη ενός κύκλου αύξησης των επιτοκίων από την FOMC (tightening cycle).

4. **Περίοδος σταθεροποίησης (2024):** Τέλος, αυτή η περίοδος εστιάζει στην επιβράδυνση του πληθωρισμού, καθώς και στη διατήρηση της ανθεκτικότητας της αγοράς εργασίας, σημαδοτώντας μια φάση προσεκτικής κανονικοποίησης.

Τα δεδομένα για την παρούσα έρευνα συλλέχθηκαν από την επίσημη ιστοσελίδα της Federal Reserve (www.federalreserve.gov). Όμως καθώς η επίσημη ιστοσελίδα με την πάροδο των ετών άλλαζε μορφή χρειάστηκε ο κώδικας να επισκεφθεί τρεις διαφορετικές σελίδες του ιστότοπου της Fed. Συγκεκριμένα:

- Για το 2018: /monetarypolicy/fomchistorical2018.htm
- Για το 2019: /monetarypolicy/fomchistorical2019.htm.
- Για 2020-2024: monetarypolicy/fomccalendars.htm

Το σύνολο των δεδομένων περιλαμβάνει πενήντα έξι (56) αρχεία FOMC Minutes, οκτώ (8) ανά έτος, τα οποία αντιστοιχούν στις οκτώ κανονικές ετήσιες συνεδριάσεις της Επιτροπής. Κάθε αρχείο αναφέρεται σε μια συγκεκριμένη ημερομηνία συνεδρίασης. Τα αρχεία αυτά μετατράπηκαν από μορφή HTML σε απλό κείμενο (.txt) και υποβλήθηκαν σε διαδικασία εκκαθάρισης και επεξεργασίας σε μορφή κατάλληλη για στατιστική ανάλυση, με την χρήση της Python.

ΠΙΝΑΚΑΣ 3.1.Β: ΠΕΡΙΓΡΑΦΙΚΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΟΛΟΥ ΔΕΔΟΜΕΝΩΝ

ΕΤΟΣ	ΑΡΧΕΙΑ	ΠΑΡΑΓΡΑΦΟΙ	ΛΕΞΕΙΣ	ΜΟ ΛΕΞΕΩΝ ΑΝΑ ΕΓΓΡΑΦΟ
2018	8	345	40.136	5.017
2019	8	346	44.286	5.536
2020	8	371	47.894	5.987
2021	8	367	46.076	5.760
2022	8	349	43.207	5.401
2023	8	317	39.329	4.916
2024	8	283	34.428	4.304
ΣΥΝΟΛΟ	**56**	**2.378**	**295.356**	**5.274**

3.2 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΕΙΜΕΝΟΥ

Μετά την συλλογή τα 56 αρχεία των FOMC Minutes υποβλήθηκαν σε μια σειρά βημάτων προεπεξεργασίας για μετατροπή σε δομημένη μορφή κατάλληλη για NLP. Ο κύριος στόχος αυτής της φάσης ήταν η μετατροπή του ακατέργαστου κειμένου, σε μορφή κατάλληλα δομημένη για υπολογιστική επεξεργασία, χωρίς όμως να αλλοιώνεται η σημασιολογική ουσία του περιεχομένου τους.

3.2.1 ΔΙΑΔΙΚΑΣΙΑ ΠΡΟΕΠΕΞΕΡΓΑΣΙΑΣ

Η διαδικασία προεπεξεργασίας ακολούθησε μια διαδοχική δομή πέντε βασικών βημάτων:

ΒΗΜΑ 1: Καθαρισμός κειμένου και εξαγωγή μεταδεδομένων

Το ακατέργαστο κείμενο, με την χρήση της βιβλιοθήκης BeautifulSoup "καθαρίστηκε" από HTML tags και διακοσμητικά στοιχεία. Έπειτα μετατράπηκε σε πεζά γράμματα και αφαιρέθηκαν ειδικοί χαρακτήρες και σημεία στίξης. Το μόνο σημείο στίξης που διατηρήθηκε ήταν η τελεία, η οποία χρησιμεύει για τον διαχωρισμό των προτάσεων και των παραγράφων.

ΒΗΜΑ 2: Tokenization και Διαχωρισμός παραγράφων

Το κείμενο διαχωρίστηκε σε παραγράφους με βάση τις τελείες και τους χαρακτήρες νέας γραμμής. Στη συνέχεια, κάθε παράγραφος διασπάστηκε σε tokens (μεμονωμένες λέξεις) χρησιμοποιώντας τον tokenizer της βιβλιοθήκες NLTK. Επιπλέον, εντοπίστηκαν και ομαδοποιήθηκαν συχνοί συνδυασμοί λέξεων σε bigrams και trigrams. Παραδείγματος χάρη, οι λέξεις Monetary policy αποτελούν ένα bigram, ενώ το Labor Market Conditions αποτελεί ένα trigram. Αυτοί οι συνδυασμοί λέξεων συχνά μεταφέρουν συγκεκριμένη οικονομική σημασία.

ΒΗΜΑ 3: Αφαίρεση Stopwords

Για να εστιάσουμε στα κύρια σημασιολογικά στοιχεία του κειμένου πραγματοποιήθηκε η αφαίρεση λέξεων που δεν προσφέρουν ουσιαστική πληροφορία. Η διαδικασία αυτή βασίστηκε στο ενσωματωμένο λεξικό stopwords της βιβλιοθήκης NLTK για την αγγλική γλώσσα. Επίσης, δημιουργήθηκε ένα συμπληρωματικό λεξικό stopwords, για καλύτερο "καθαρισμό" που περιλάμβανε τις παρακάτω 3 κατηγορίες:

1. Θεσμικούς και Οργανωτικούς όρους: Δηλαδή λέξεις που αναφέρονται στη λειτουργία και στη δομή της FOMC, οι οποίες εμφανίζονται συστημικά σε όλα τα πρακτικά αλλά φέρουν περιορισμένη διακριτική ισχύ για τη θεματική ανάλυση. Κάποια παραδείγματα είναι "meeting", "participant", "member", "vote".

2. Γενικούς οικονομικούς όρους: Λέξεις που εμφανίζονται συχνά και ενώ ανήκουν στο οικονομικό λεξιλόγιο, δεν συμβάλλουν στην διαφοροποίηση των κειμένων. Κάποια παραδείγματα: "economic", "policy", "rate", "market".

3. Διαδικαστικές λέξεις: Λέξεις που περιγράφουν την διαδικασία συζήτησης ή τον τρόπο παρουσίασης μιας άποψης, χωρίς όμως να περιέχουν ουσιαστική πληροφορία. Κάποια παραδείγματα: "noted", "commented", "suggested", "observed".

Αυτό το βήμα μας πρόσφερε "μείωση θορύβου" στα δεδομένα καθώς αφαιρέθηκαν ≈22,2% λέξεις που δεν μας πρόσφεραν ουσιαστικές πληροφορίες για την ανάλυση.

ΒΗΜΑ 4: Lemmatization

Για να επιτευχθεί λεξιλογική συνοχή και να βελτιωθεί η ποιότητα των γλωσσικών αναπαραστάσεων, εφαρμόστηκε η διαδικασία lemmatization χρησιμοποιώντας το WordNetLemmatizer της βιβλιοθήκης NLTK. Αυτό έχει ως αποτέλεσμα να κρατάμε την ρίζα των λέξεων και να ομαδοποιούμε τις λέξεις με κοινές ρίζες.

ΒΗΜΑ 5: Ομαδοποίηση και Ταξινόμηση

Τα επεξεργασμένα δεδομένα οργανώθηκαν σε ένα πλήρως δομημένο Data Frame με τις ακόλουθες στήλες:

- Document_id: Μοναδικό αναγνωριστικό εγγράφου
- Date: Ημερομηνία συνεδρίασης
- Year: Έτος για χρονική ομαδοποίηση
- Paragraph_id: Αναγνωριστικό παραγράφου εντός εγγράφου
- Raw_text: Αρχικό ακατέργαστο κείμενο
- Processed_text: Τελικό επεξεργασμένο κείμενο
- Word_count: Αριθμός λέξεων μετά την προεπεξεργασία
- Token_count: Αριθμός token (συμπεριλαμβανομένων n-grams)

Μετά τον διαχωρισμό σε παραγράφους και την προεπεξεργασία, ο μέσος όρος λέξεων ανά παράγραφο μειώθηκε κατά ≈25,5% (από 124,4 σε 92,7 λέξεις), αυτό οφείλεται στην αφαίρεση των stopwords και της lemmatization.

3.2.2 ΑΠΟΘΗΚΕΥΣΗ ΚΑΙ ΔΙΑΧΕΙΡΙΣΗ ΔΕΔΟΜΕΝΩΝ

Τα δεδομένα αποθηκεύτηκαν σε δύο βασικά αρχεία, κάθε ένα με διαφορετικό σκοπό:

1. fomc_sample.pkl: Το αρχείο αυτό περιέχει ένα δείγμα των μη-επεξεργασμένων κειμένων με πλήρη διατήρηση της αρχικής δομής, HTML tags και μεταδεδομένων. Αυτό το αρχείο χρησιμεύει ως επαλήθευση και ανάκτηση των αρχικών πληροφοριών ανά πάσα στιγμή, χωρίς να χρειάζεται να ξανά γίνει η διαδικασία του κατεβάσματος, η οποία είναι χρονοβόρα.

2. Fomc_processed.pkl: Ενώ αυτό το αρχείο περιέχει τα πλήρως επεξεργασμένα δεδομένα σε μορφή κατάλληλη για ανάλυση. Τα πλεονεκτήματα του είναι ότι είναι τουλάχιστον 40-60% μικρότερο σε μέγεθος αρχείο, αυτό έχει ως αποτέλεσμα να είναι πιο αποδοτικό και να φορτώνει με μεγαλύτερη ταχύτητα.

3.3 ΑΝΑΛΥΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (SENTIMENT ANALYSIS)

Η Ανάλυση Συναισθήματος αποτελεί τον βασικό πυρήνα της μεθοδολογίας για αυτή την έρευνα. Έχει ως στόχο την ποσοτικοποίηση και την ανάλυση του συναισθηματικού τόνου των FOMC Minutes. Στη παρούσα έρευνα χρησιμοποιήθηκαν δύο προσεγγίσεις: Μια λεξικολογική και μια βασισμένη σε μοντέλο μηχανικής μάθησης. Αυτό μας εξασφαλίζει την αξιοπιστία, αλλά και την ολοκληρωμένη αντίληψη των διάφορων πτυχών του συναισθήματος.

3.3.1 ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Η πρώτη προσέγγιση βασίστηκε στο Loughran-McDonald Financial Sentiment Dictionary.. Το λεξικό αυτό προσεγγίζει το συναίσθημα ταξινομώντας τις λέξεις σε επτά βασικές κατηγορίες: Αρνητικό (Negative), Θετικό (Positive), Νομικό (Litigious), Αβεβαιότητα (Uncertainty), Περιοριστικό (Constraining), Ισχυρή Τροπικότητα (Strong Modal)και Ασθενής Τροπικότητα (Weak Modal).

Αφού χωριστούν οι λέξεις στις παραπάνω κατηγορίες, για τον υπολογισμό του Sentiment Score(S_i) σε κάθε παράγραφο i χρησιμοποιείται ο εξής τύπος:

$$S_i = \frac{P_i - N_i}{Ti}$$

Όπου:

P_i= Αριθμός θετικών λέξεων στην i

N_i= Αριθμός αρνητικών λέξεων στην i

T_i= Συνολικός αριθμός λέξεων στην i

Το αποτέλεσμα του Si είναι ένας αριθμός στο κλειστό διάστημα [-1,1]. Μια τιμή κοντά στο -1 υποδηλώνει κυριαρχία αρνητικού συναισθήματος, μια τιμή κοντά στο 1 υποδηλώνει έντονα θετικό συναίσθημα, ενώ όταν η τιμή ισούται με το 0 υποδεικνύει είτε ισορροπία μεταξύ θετικών και αρνητικών λέξεων, είτε πλήρη απουσία τους.

Το λεξικό Loughran-McDonald επιλέχθηκε επειδή είναι εξειδικευμένο στον οικονομικό τομέα, έχει ευρεία ακαδημαϊκή αποδοχή και χαρακτηρίζεται από απόλυτη διαφάνεια και επαναληψιμότητα. Ωστόσο, η μέθοδος αυτή έχει και περιορισμούς. Οι κύριες αδυναμίες της είναι η έλλειψη κατανόησης του συμφραζόμενου (π.χ., η λέξη "strong" είναι θετική στο "strong growth" αλλά αρνητική στο "strong headwinds") και η έλλειψη αναγνώρισης της μεταφορικής γλώσσας (π.χ. "overheating economy").

3.3.2 ΠΡΟΣΕΓΓΙΣΗ ΜΟΝΤΕΛΟΥ ΜΗΧΑΝΙΚΗΣ ΜΑΘΗΣΗΣ

Για να αντιμετωπιστούν οι περιορισμοί της λεξιλογικής προσέγγισης που αναφέρθηκαν παραπάνω, χρησιμοποιήθηκε το γλωσσικό μοντέλο FinBERT (Financial Bidirectional Encoder Representations from Transformers) και είναι ένα προηγμένο, προ-εκπαιδευμένο γλωσσικό μοντέλο, το οποίο έχει εκπαιδευτεί ειδικά σε ένα τεράστιο corpus χρηματοοικονομικών κειμένων.

Σε αντίθεση με το λεξικό, το FinBERT έχει την δυνατότητα να καταλαβαίνει την σημασία των λέξεων εντός συμφραζομένου. Όμως έχει τον περιορισμό να αναλύει έως 512 tokens ανά είσοδο. Για να αντιμετωπιστεί αυτό, το κείμενο διαχωρίστηκε σε λογικά τμήματα (chunks) διατηρώντας τη συνοχή των προτάσεων.

Τα αποτελέσματα που εξάγουμε από αυτή την προσέγγιση είναι μια τριμερής ταξινόμηση positive, negative, neutral, συνοδευόμενη από έναν βαθμό βεβαιότητας (confidence score) για κάθε ταξινόμηση. Αυτό του επιτρέπει να διακρίνει "αποχρώσεις", να αναγνωρίζει μεταφορές και να εκτιμάει το "συναίσθημα" ολόκληρων των προτάσεων και παραγράφων, λαμβάνοντας υπόψη την δομή και την σύνταξη.

Για να υπολογίσουμε το sentiment score για μια παράγραφο i, χρησιμοποιούμε τον μέσο όρο του καθαρού συναισθήματος(Net Sentiment) για όλα τα τμήματα(chunks) που την αποτελούν.

Το καθαρό συναίσθημα για κάθε τμήμα j υπολογίζεται από τον εξής τύπο:

$$\text{Net Sentiment}_j = P_j(\text{positive}) - P_j(\text{negative})$$

Όπου:

P_j(positive) = Πιθανότητα το τμήμα j να είναι θετικό

P_j(negative) = Πιθανότητα το τμήμα j να είναι αρνητικό

Αυτή η τιμή κυμαίνεται σε κλειστό διάστημα [-1,1] , όπου οι τιμές κοντά στο 1 υποδηλώνουν θετική πιθανότητα, στο -1 αρνητική , ενώ οι τιμές γύρω στο 0 είτε ουδέτερο ή ανάμεικτο συναίσθημα.

Ακολούθως , για να υπολογιστεί το sentiment score για την παράγραφο i με τμήματα k υπολογίζεται ως:

$$\text{Sentiment Score}_i = \frac{1}{k}\sum_{j=1}^{k} \text{Net Sentiment}_j$$

Δηλαδή όπως αναφέρθηκε και προηγουμένως είναι ο μέρος όρος των καθαρών συναισθημάτων όλων των τμημάτων.

3.4 ΘΕΜΑΤΙΚΗ ΑΝΑΛΥΣΗ (LATENT DIRICHLET ALLOCATION-LDA)

Στην παρούσα έρευνα χρησιμοποιήθηκε το μοντέλο LDA για να προσδιορισθούν οι κύριοι θεματικοί άξονες που απασχολούν την FOMC για την χρονική περίοδο 2018-2024 και να ποσοτικοποιηθεί η σημασία των διαφόρων οικονομικών εννοιών ανά περίοδο. Επιπλέον, αυτό το μοντέλο βοηθάει στην έρευνα και την ανάλυση των μεταβαλλόμενων προτεραιοτήτων της Επιτροπής στην πάροδο των ετών, αλλά και στον συνδυασμό με τα

αποτελέσματα της Ανάλυσης Συναισθήματος (Sentiment Analysis) για μια πιο ολοκληρωμένη ερμηνεία.

3.4.1 ΥΛΟΠΟΙΗΣΗ LDA

Για την επίτευξη μιας ολοκληρωμένης ανάλυσης εφαρμόστηκε το μοντέλο της LDA σε δύο διαφορετικά επίπεδα. Αρχικά, έγινε ανάλυση σε επίπεδο παραγράφου και έπειτα σε επίπεδο έτους. Στην ανάλυση σε επίπεδο παραγράφου εξετάζεται κάθε παράγραφος ξεχωριστά, ώστε να κατανοηθούν οι λεπτές διαφοροποιήσεις στις συζητήσεις. Αυτό βοηθάει στον εντοπισμό των βασικών και σταθερών θεμάτων που εμφανίζονται σε όλα τα έτη, ανεξάρτητα από τις συνθήκες κάθε εποχής. Αντίθετα, η ανάλυση στο επίπεδο του έτους έχει ως στόχο την ομαδοποίηση των δεδομένων ανά χρονιά. Αυτή η προσέγγιση επιτρέπει να παρατηρηθεί πως εξελίσσονται και μεταβάλλονται οι συζητήσεις και συγκεκριμένα πως ανταποκρίνεται η FOMC σε σημαντικά γεγονότα, όπως αυτό της πανδημίας COVID-19.

3.4.2 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ

Η πρώτη προσέγγιση που χρησιμοποιήθηκε εστιάζει στο επίπεδο της παραγράφου. Αυτή η επιλογή οφείλεται στην παρατήρηση ότι κάθε παράγραφος στα πρακτικά της FOMC λειτουργεί ως ένα μικρό "σημειολογικό παράθυρο" σε μία συγκεκριμένη σκέψη ή επιχείρημα. Μέσα από την ανάλυση των μεμονωμένων παραγράφων, μπορούν να εντοπιστούν λεπτές διαφοροποιήσεις στις συζητήσεις. Ουσιαστικά η ανάλυση LDA ανά παράγραφο λειτουργεί ως ένα "μικροσκόπιο", το οποίο προσφέρει την δυνατότητα να εντοπιστούν τα βασικά και σταθερά θέματα που διατρέχουν όλα τα έτη ανεξάρτητα από τις ειδικότερες συνθήκες κάθε περιόδου.

3.4.2.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Η διαδικασία ξεκίνησε με τη φόρτωση των προεπεξεργασμένων δεδομένων από το αρχείο fomc_processed.pkl. Η προετοιμασία των δεδομένων για το LDA περιλάμβανε επιπλέον φιλτράρισμα. Συγκεκριμένα, από το σύνολο των παραγράφων, διατηρήθηκαν μόνο εκείνες με τουλάχιστον 5 λέξεις και μόνο λέξεις συγκεκριμένων γραμματικών κατηγοριών (όπως ουσιαστικά, ρήματα,

επίθετα, κύρια ονόματα και συντομογραφίες) για να διατηρηθεί η σημασιολογική πυκνότητα. Επίσης, πραγματοποιήθηκε αφαίρεση λέξεων μήκους μικρότερων των 2 χαρακτήρων και καθαρισμός από σημεία στίξης και ειδικούς χαρακτήρες. Έτσι διασφαλίστηκε ότι οι παράγραφοι που θα αναλυθούν θα περιέχουν επαρκή πληροφορία, ώστε να παρέχουν ουσιαστικά αποτελέσματα.

3.4.2.2 ΔΗΜΙΟΥΡΓΙΑ ΛΕΞΙΛΟΓΙΟΥ

Από το συνολικό corpus, που περιλάμβανε 8.542 διαφορετικές λέξεις, δημιουργήθηκε ένα ειδικό λεξιλόγιο για το LDA. Για την δημιουργία του εφαρμόστηκαν δύο βασικά φίλτρα:

- no_below=3: Διαγραφή λέξεων που εμφανίζονται σε λιγότερες από τρείς παραγράφους, καθώς θεωρήθηκαν πολύ σπάνιες για να έχουν στατιστική σημασία.
- no_above=0.7: Διαγραφή λέξεων που εμφανίζονται σε περισσότερο από το 70% των παραγράφων. Καθώς εμφανίζονταν πολύ συχνά και πιθανώς είναι μη διακριτές.

Μετά το φιλτράρισμα, το λεξιλόγιο περιείχε 1.245 λέξεις. Στη συνέχεια, κάθε παράγραφος μετατράπηκε σε μια διανυσματική αναπαράσταση (bag of words) που καταγράφει τη συχνότητα εμφάνισης κάθε λέξης του λεξιλογίου σε αυτήν, δημιουργώντας την είσοδο για τον αλγόριθμο LDA.

3.4.2.3 ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ LDA

Το πιο σημαντικό βήμα της ανάλυσης ήταν η εκπαίδευση του ίδιου του μοντέλου LDA. Για να επιλεχθεί ο αριθμός topics, υιοθετήθηκε μια προσαρμοστική προσέγγιση που βασίζεται στο μέγεθος του dataset.

ΠΙΝΑΚΑΣ 3.4.Α: ΚΡΙΤΗΡΙΑ ΓΙΑ ΑΡΙΘΜΟ TOPICS ΣΤΗΝ ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ

ΑΡΙΘΜΟΣ ΠΑΡΑΓΡΑΦΩΝ	ΑΡΙΘΜΟΣ TOPICS
<20	2
20-50	3

50-100	4
>100	5

Δεδομένου ότι το dataset περιείχε 1.876 παραγράφους, επιλέχθηκαν 5 topics για την ανάλυση.

Το μοντέλο εκπαιδεύτηκε με τις παραμέτρους 50 επαναλήψεων (iterations) ανά έγγραφο και 10 πλήρη περάσματα (passes) του συνόλου δεδομένων για να διασφαλιστεί η πλήρης σύγκλιση και η σταθερότητα των εκτιμήσεων.

3.4.3 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ

Η δεύτερη προσέγγιση μεταφέρει την ανάλυση σε ένα πιο μακροσκοπικό επίπεδο. Σε αυτή την προσέγγιση ομαδοποιήθηκαν όλα τα δεδομένα ανά έτος και εφαρμόζεται LDA ξεχωριστά σε κάθε έτος. Με στόχο την ανάλυση της χρονικής εξέλιξης και της μετατόπισης των προτεραιοτήτων της της FOMC ως απάντηση σε σημαντικά γεγονότα (πχ. πανδημία COVID-19, πληθωριστική κρίση).

3.4.3.1 ΕΠΕΞΕΡΓΑΣΙΑ ΔΕΔΟΜΕΝΩΝ

Η διαδικασία ξεκίνησε με τη φόρτωση του αρχείου fomc_sample.pkl, το οποίο περιέχει τις αρχικές, μη-επεξεργασμένες εκδόσεις των πρακτικών. Χρησιμοποιήθηκε ώστε η διαχρονική ανάλυση να γίνει με βάση την πιο πλήρη και αμετάβλητη έκδοση των δεδομένων.

Για να επιτευχθεί η ομαδοποίηση όλων των παραγράφων ανά έτος, χρειάστηκε να εξαχθούν τα έτη από κάθε ημερομηνία συνεδρίασης, σε μορφή YYYY-MM-DD. Έπειτα, ομαδοποιήθηκαν όλες οι παράγραφοι που ανήκαν στο ίδιο έτος και έτσι δημιουργήθηκαν ετήσια σύνολα δεδομένων.

ΠΙΝΑΚΑΣ 3.4.Β: ΣΤΑΤΙΣΤΙΚΑ ΟΜΑΔΟΠΟΙΗΣΗΣ ΠΑΡΑΓΡΑΦΩΝ

ΕΤΟΣ	ΑΡΙΘΜΟΣ ΣΥΝΕΔΡΙΑΣΕΩΝ	ΠΑΡΑΓΡΑΦΟΙ
2018	8	345
2019	8	346
2020	8	371
2021	8	367
2022	8	349
2023	8	317
2024	8	283
ΣΥΝΟΛΟ	**56**	**2378**

3.4.3.2 ΠΡΟΕΠΕΞΕΡΓΑΣΙΑ ΚΕΙΜΕΝΟΥ ΓΙΑ ΚΑΘΕ ΕΤΟΣ

Για κάθε έτος χωριστά, τα δεδομένα υποβλήθηκαν στα εξής βήματα για προεπεξεργασία. Αρχικά, μετατράπηκε το κείμενο σε πεζά γράμματα, αφαιρέθηκαν τα σημεία στίξης και τα stopwords. Διατηρήθηκαν λέξεις με περισσότερους από δύο χαρακτήρες και παράγραφοι με τουλάχιστον πέντε λέξεις. Στα έτη όπου είχαν λιγότερες παραγράφους, χρησιμοποιήθηκαν πιο χαλαρά κριτήρια και συγκεκριμένα διατηρήθηκαν παράγραφοι με τουλάχιστον 3 λέξεις, ώστε να διασφαλιστεί επαρκής όγκος δεδομένων για την ανάλυση.

3.4.3.3 ΕΚΠΑΙΔΕΥΣΗ ΜΟΝΤΕΛΟΥ LDA

Η διαδικασία εκπαίδευσης LDA εφαρμόστηκε ξεχωριστά για κάθε έτος. Υιοθετήθηκε μια προσαρμοστική στρατηγική που λαμβάνει υπόψη το μέγεθος του dataset κάθε έτους:

Στα έτη με λιγότερα δεδομένα (πχ. 2018) εφαρμόστηκαν πιο χαλαρά κριτήρια στο λεξιλόγιο no_below=2, no_above=0.9 και ο αριθμός των topics ήταν 3-4, ανάλογα τον όγκο των δεδομένων. Στην εκπαίδευση δόθηκε έμφαση στη σταθερότητα με περισσότερες επαναλήψεις. Ενώ αντίθετα στα έτη με μεγάλο όγκο δεδομένων (πχ. 2020-2022)εφαρμόστηκαν πιο αυστηρά κριτήρια για το λεξιλόγιο no_below=3, no_above=0.7 και ο αριθμός των

topics ήταν 5. Στην εκπαίδευση δόθηκε έμφαση στην βελτιστοποίηση για ευρεία θεματική κάλυψη.

ΠΙΝΑΚΑΣ 3.4.Γ: ΚΡΙΤΗΡΙΑ ΓΙΑ ΑΡΙΘΜΟ TOPICS ΣΤΗΝ ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ

ΑΡΙΘΜΟΣ ΠΑΡΑΓΡΑΦΩΝ	ΑΡΙΘΜΟΣ TOPICS
<30	3
30-60	4
>100	5

Σε όλα τα έτη χρησιμοποιήθηκαν οι ίδιες βασικές παράμετροι, ώστε να είναι συγκρίσιμα. Έτσι έγιναν 50 επαναλήψεις ανά έγγραφο και 10 ολόκληρα περάσματα του dataset.

3.5 ΣΥΓΚΡΙΤΙΚΗ ΑΞΙΟΛΟΓΗΣΗ ΜΕΘΟΔΟΛΟΓΙΩΝ

Ο συνδυασμός αυτών των τεσσάρων μεθόδων δεν είναι τυχαίος, σχεδιάστηκε με σκοπό η κάθε μέθοδος να συμπληρώνει τις άλλες, δημιουργώντας ένα ισχυρό αναλυτικό πλαίσιο που ξεπερνάει τους περιορισμούς οποιασδήποτε μεμονωμένης προσέγγισης.

Ο πρώτος συνδυασμός LDA και Lexicon επιτρέπει να απαντηθεί το ερώτημα: «Τι συζητάνε και πώς το συζητάνε;». Η LDA δίνει έμφαση στα κύρια θέματα των συζητήσεων, ενώ το lexicon αποκαλύπτει τον συναισθηματικό τόνο για το κάθε ένα. Για παράδειγμα, μπορεί να αναγνωρίσει αν τα topics που αφορούν τον πληθωρισμό έχουν αρνητικό ή θετικό συναίσθημα.

Ο δεύτερος συνδυασμός LDA και FinBERT προσφέρει μια πιο εμβάθυνση, καθώς επιτρέπει να απαντηθεί το ερώτημα «Ποια είναι η δομή των συζητήσεων και ποιες είναι οι "αποχρώσεις" των συναισθημάτων;». Το LDA αποκαλύπτει τις θεματικές κατηγορίες των κειμένων, ενώ το FinBERT επιτρέπει την κατανόηση των ανάμεικτων συναισθημάτων καθώς και τις πολύπλοκες σημασιολογικές σχέσεις.

Τέλος, ο τρίτος συνδυασμός ανάλυσης στο επίπεδο της παραγράφου και στο επίπεδο του έτους, δημιουργεί μια χρονική διάσταση στην ανάλυση και απαντά στο ερώτημα: «Πως εξελίσσονται οι ιδέες από την μια παράγραφο στην άλλη και από το ένα έτος στο άλλο;». Αυτός ο συνδυασμός

επιτρέπει την παρακολούθηση της εξέλιξης συγκεκριμένων εννοιών και τον εντοπισμό των στοιχείων που μετατρέπονται σε προτεραιότητα της Επιτροπής με την πάροδο του χρόνου.

Μετά τη λεπτομερή περιγραφή της μεθοδολογίας που ακολουθήθηκε, έχουμε όλα τα απαραίτητα εργαλεία και πλαίσια για να προχωρήσουμε στην παρουσίαση και ανάλυση των αποτελεσμάτων. Το Κεφάλαιο 3 αποτύπωσε πώς θα απαντήσουμε στα ερευνητικά μας ερωτήματα, ενώ το Κεφάλαιο 4 έχει ως στόχο να παρουσιάσει τα αποτελέσματα και τις απαντήσεις. Η πολύ-μεθοδική προσέγγιση που περιγράφηκε σε αυτό το Κεφάλαιο θα μας επιτρέψει να προσεγγίσουμε τα αποτελέσματα από διάφορες οπτικές γωνίες.

3.6 ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ

Η προηγούμενη ανάλυση (ενότητες 3.1 έως 3.5) επικεντρώθηκε αποκλειστικά στην εξαγωγή πληροφοριών από το κείμενο των FOMC Minutes (Sentiment, topics). Ωστόσο, για να αξιολογηθεί η ευρύτερη πολιτική διαδικασία, είναι απαραίτητο να τοποθετηθούν αυτά τα γλωσσικά ευρήματα στο χρονικό και πολιτικό πλαίσιο στο οποίο παραχθήκαν. Η παρούσα ενότητα περιγράφει τη μεθοδολογία συλλογής και επεξεργασίας των συμπληρωματικών δεδομένων που απαιτήθηκαν για αυτή τη διευρυμένη ανάλυση, η οποία θα παρουσιαστεί στο Κεφάλαιο 4.4.

Επομένως, για τη διενέργεια της συγκριτικής ανάλυσης μεταξύ γλωσσικών ευρημάτων και πολιτικών παρεμβάσεων στο Κεφάλαιο 4.4, συλλέχθηκαν και επεξεργάστηκαν δεδομένα από συμπληρωματικές πηγές, πέραν των FOMC Minutes:

1. Δεδομένα Πολιτικών Παρεμβάσεων: Καταρτίστηκε ένα λεπτομερές χρονολόγιο των 15 πιο σημαντικών νομισματικών (π.χ. μείωση επιτοκίων, απεριόριστη QE) και δημοσιονομικών παρεμβάσεων (π.χ. CARES Act, American Rescue Plan) για την περίοδο 2020-2021. Τα δεδομένα συλλέχθηκαν από επίσημες πηγές όπως η ιστοσελίδα της Federal Reserve και του U.S. Department of the Treasury.

2. Οικονομικοί Δείκτες: Για τη στατιστική ανάλυση συσχέτισης , χρησιμοποιήθηκαν δημόσια διαθέσιμοι εβδομαδιαίοι και μηνιαίοι

δείκτες, κυρίως σχετικά με την αγορά εργασίας (π.χ. εβδομαδιαίες αιτήσεις ανεργίας). Οι πηγές περιλαμβάνουν το U.S. Bureau of Labor Statistics (BLS).

3. Επεξεργασία και Ομαδοποίηση: Τα δεδομένα από όλες τις πηγές (πολιτικά γεγονότα, οικονομικοί δείκτες, γλωσσικά σκορ από τα FOMC Minutes) ομαδοποιήθηκαν και ευθυγραμμίστηκαν σε μηνιαία και τριμηνιαία βάση. Αυτή η χρονική συγχρονισμένη δομή επέτρεψε την άμεση σύγκριση και αντιστοίχιση μεταξύ: της χρονικής στιγμής μιας πολιτικής ανακοίνωσης, της μεταβολής στον συναισθηματικό τόνο των πρακτικών της FOMC και της εξέλιξης βασικών οικονομικών δεικτών.

4. Μέθοδοι Στατιστικής Ανάλυσης: Για τον υπολογισμό της σχέσης μεταξύ του συναισθήματος και των οικονομικών δεικτών (π.χ. ανεργία) χρησιμοποιήθηκε ο συντελεστής συσχέτισης Pearson (r), ο οποίος μετρά την ισχύ και την κατεύθυνση μιας γραμμικής σχέσης μεταξύ δύο συνεχών μεταβλητών.

Σκοπός αυτής της ανάλυσης είναι να μετατραπεί η ποιοτική και γλωσσική ανάλυση των προηγούμενων ενοτήτων σε ένα πλαίσιο ποσοτικής πολιτικής αξιολόγησης, επιτρέποντας τη κατανόηση της διαδικασίας λήψης αποφάσεων κατά την κρίση.

ΚΕΦΑΛΑΙΟ 4: ΑΝΑΛΥΣΗ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Το παρόν κεφάλαιο αποσκοπεί στην παρουσίαση και στην κριτική ερμηνεία των ευρημάτων που προέκυψαν από την πολύ-μεθοδική ανάλυση που περιγράφηκε στο Κεφάλαιο 3. Σκοπός είναι να απαντηθούν τα ερευνητικά ερωτήματα σχετικά με τον συναισθηματικό τόνο, τις θεματικές προτεραιότητες και τη χρονική εξέλιξη στα πρακτικά της FOMC για την κρίσιμη περίοδο 2018-2024. Τα αποτελέσματα θα παρουσιαστούν με τη σειρά που ακολουθήθηκε στη μεθοδολογία. Αρχικά, θα παρουσιαστεί η ανάλυση του συναισθήματος με τις δύο προσεγγίσεις: Lexicon και FinBERT και στη συνέχεια η θεματική ανάλυση LDA σε επίπεδο παραγράφου και έτους. Η ερμηνεία θα γίνει σε συνεχή αναφορά με το ιστορικό και οικονομικό πλαίσιο κάθε υποπεριόδου όπως ορίστηκε στον Πίνακα 3.1.Α,

επιβεβαιώνοντας έτσι τη δυναμική σχέση μεταξύ γλώσσας, πολιτικής και οικονομικών συνθηκών.

4.1 ΑΠΟΤΕΛΕΣΜΑΤΑ ΑΝΑΛΥΣΗΣ ΣΥΝΑΙΣΘΗΜΑΤΟΣ

Σε αυτή την ενότητα παρουσιάζονται τα αποτελέσματα της ποσοτικοποίησης του συναισθηματικού τόνου των πρακτικών, με χρήση τόσο της λεξιλογικής προσέγγισης (Loughran-McDonald) όσο και του εξειδικευμένου μοντέλου FinBERT.

4.1.1 ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

Η εφαρμογή του λεξικού Loughran-McDonald στις 2.378 παραγράφους του corpus αποκάλυψε ότι ο γλωσσικός τόνος των πρακτικών είναι σταθερά προσεκτικός και τείνει προς αρνητικός για την περίοδο 2018-2024. Οι διακυμάνσεις αυτού του τόνου αντανακλούν άμεσα τις οικονομικές προκλήσεις κάθε περιόδου.

Βάσει της μεθοδολογίας του λεξικού, οι λέξεις ταξινομήθηκαν σε επτά κατηγορίες, με τις κατηγορίες "Αρνητικό"(Negative) και "Νομικό"(Litigious) να καταλαμβάνουν το μεγαλύτερο ποσοστό. Ενώ ο αριθμός των Θετικών όρων(Positive) είναι σχετικά μικρός, όπως θα παρατηρήσουμε στον Πίνακα 4.1.Α που ακολουθεί.

ΠΙΝΑΚΑΣ 4.1.Α: ΚΑΤΑΝΟΜΗ ΛΕΞΕΩΝ ΑΝΑ ΚΑΤΗΓΟΡΙΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (ΛΕΞΙΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ)

ΚΑΤΗΓΟΡΙΑ	ΑΡΙΘΜΟΣ ΛΕΞΕΩΝ	ΠΟΣΟΣΤΑ ΣΥΝΟΛΟΥ
Negative	2.345	45.2%
Litigious	903	17.4%
Uncertainty	297	5.7%
Positive	347	6.7%
Constraining	184	3.5%

Weak Modal	27	0.5%
Strong Modal	19	0.4%

Ο μέσος όρος του Sentiment Score για ολόκληρη την περίοδο ήταν -0,0171, δηλαδή ελαφρώς αρνητικός. Ωστόσο, η εξέλιξη ανά έτος αποκαλύπτει σημαντικές τάσεις που σχετίζονται άμεσα με τα μακροοικονομικά γεγονότα.

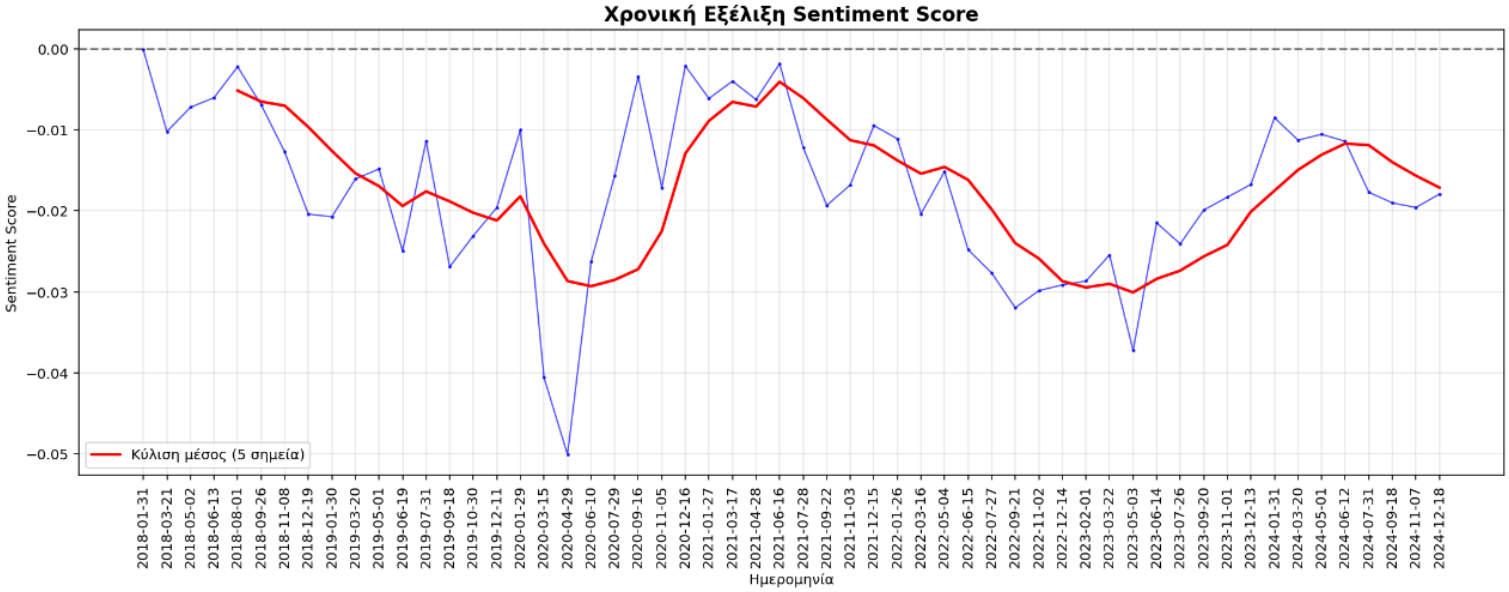
ΠΙΝΑΚΑΣ 4.1.Β: Ετήσια Στατιστικά Συναισθήματος(LEXICON)

ΕΤΟΣ	ΜΕΣΟ SCORE	% ΘΕΤΙΚΩΝ ΛΕΞΕΩΝ	% ΑΡΝΗΤΙΚΩΝ ΛΕΞΕΩΝ	ΚΑΤΑΣΤΑΣΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ
2018	-0,0082	2,14%	3,08%	Ουδέτερο
2019	-0,0197	1,87%	3,91%	Αρνητικό
2020	-0,0205	2,14%	4,36%	Αρνητικό
2021	-0,0095	2,79%	3,64%	Ουδέτερο
2022	-0,0233	1,91%	4,41%	Αρνητικό
2023	-0,0242	1,62%	4,24%	Αρνητικό
2024	-0,0143	2,19%	3,73%	Αρνητικό

Από τον Πίνακα 4.1.Β παρατηρούνται τρείς ευδιάκριτες κορυφώσεις αρνητικού τόνου στα πρακτικά της FOMC:

1. **2020** (score: -0,0205): Αντανακλά το σοκ της πανδημίας COVID-19 και την απότομη οικονομική συρρίκνωση.
2. **2022** (score: -0,0233): Αντιστοιχεί στην εκτόξευση του πληθωρισμού και την έναρξη του κύκλου αυστηροποίησης της νομισματικής πολιτικής (tightening cycle).
3. **2023** (score: -0,0242): Σχετίζεται με τη συνεχιζόμενη πίεση υψηλού πληθωρισμού και την τραπεζική αναταραχή με τη πτώση Silicon Valley Bank τον Μάρτιο, καταγράφοντας το χαμηλότερο sentiment score της περιόδου.

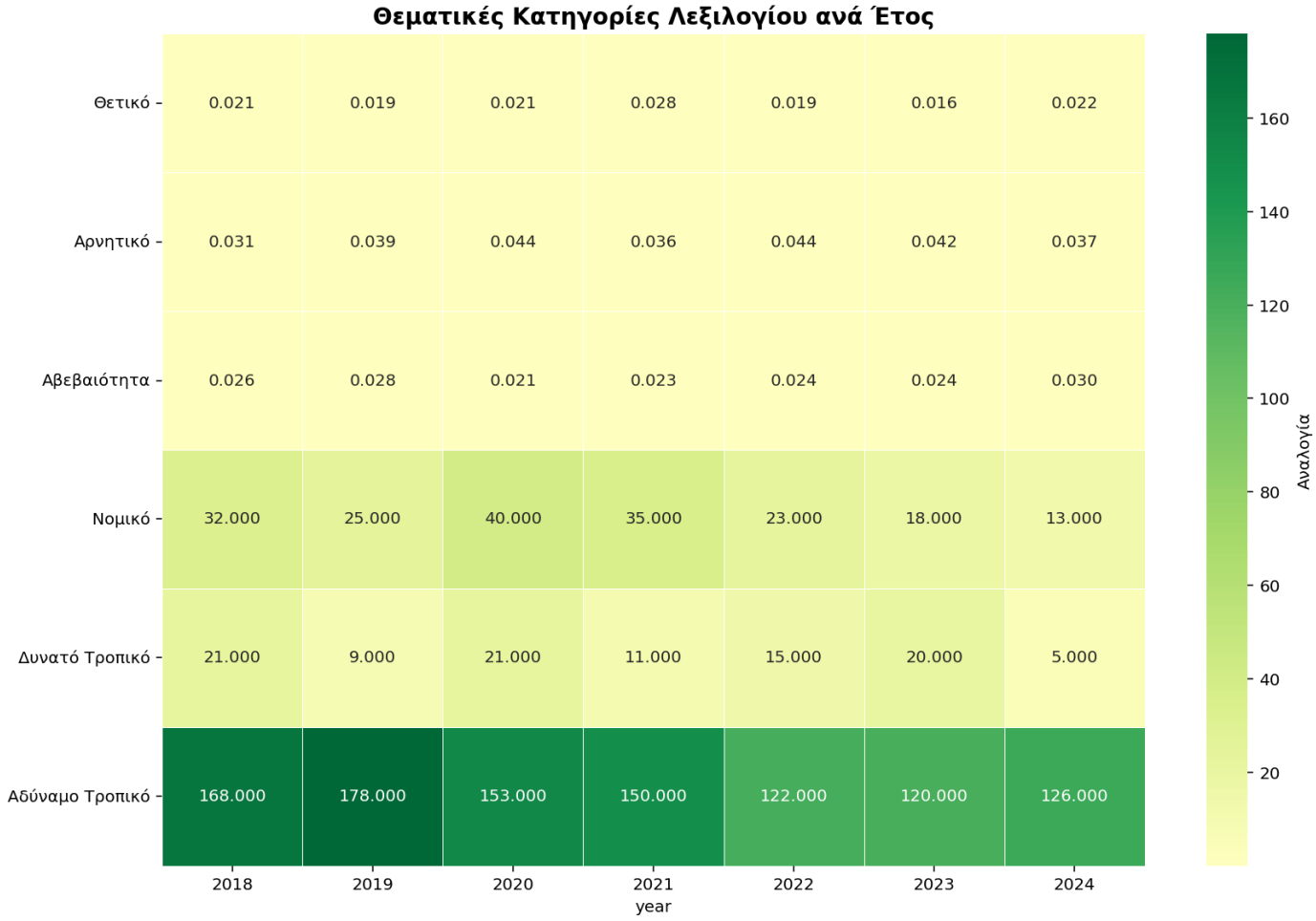
ΔΙΑΓΡΑΜΜΑ 4.1.Γ: ΧΡΟΝΙΚΗ ΕΞΕΛΙΞΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ 2018-2024 (ΛΕΞΙΚΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ)



Το διάγραμμα 4.1.Γ παρουσιάζει τη χρονική εξέλιξη του μέσου Sentiment Score για κάθε συνεδρίαση μεταξύ των ετών 2018 με 2024. Παρατηρείται μια σταθερή κάθοδος προς τις αρνητικές τιμές κατά τα έτη 2019 και 2020, φτάνοντας στο χαμηλότερο σημείο κατά την περίοδο της πανδημίας COVID-19. Έπειτα, ακολουθεί μια μικρή αλλά αισθητή ανάκαμψη το 2021, η οποία συμπίπτει με την οικονομική ανάπτυξη που σημειώθηκε με την άφιξη των εμβολίων. Ωστόσο, μια νέα και ακόμα πιο έντονη κάθοδος δημιουργείται τα έτη 2022 και 2023, η οποία αντιστοιχεί με την κορύφωση του πληθωρισμού και την έναρξη αυστηροποίησης της νομισματικής πολιτικής, καθώς και με τις αναταραχές του τραπεζικού συστήματος που σημειώθηκαν. Τέλος, το 2024 σημειώνεται βελτίωση, υποδηλώνοντας μια πρώιμη ανακούφιση καθώς ο πληθωρισμός αρχίζει να επιβραδύνεται, αν και ο συνολικός τόνος παραμένει προσεκτικός, αντικατοπτρίζοντας την επιφύλαξη της Επιτροπής. Παρατηρώντας τις ακραίες τιμές μπορεί να διαπιστωθεί ότι στις 29 Ιουλίου του 2020 σημειώνεται η πιο αρνητική παράγραφος, καθώς απεικονίζεται το απόγειο της οικονομικής αναταραχής από την πανδημία. Ο υψηλός αριθμός των αρνητικών λέξεων πιθανώς αντιστοιχεί σε περιγραφές της ραγδαίας συρρίκνωσης, των υψηλών επιπέδων ανεργίας και των σημαντικών κινδύνων που αντιμετώπιζε η οικονομία. Ενώ στις 16 Δεκεμβρίου 2020 εντοπίζεται η πιο θετική παράγραφος. Αυτή η ημερομηνία συμπίπτει χρονικά με την έγκριση των πρώτων εμβολίων κατά

την περίοδο του COVID-19 στις ΗΠΑ και την έναρξη της μαζικής προσπάθειας για εμβολιασμό. Είναι πιθανό ότι οι συζητήσεις της Επιτροπής είναι πιο αισιόδοξες καθώς εμφανίζεται μια αρχική ανακούφιση και βελτιωμένες προοπτικές για οικονομική ανάπτυξη σε μια ιδιαίτερα δύσκολη χρονιά.

ΔΙΑΓΡΑΜΜΑ 4.1.Δ: ΘΕΜΑΤΙΚΕΣ ΚΑΤΗΓΟΡΙΕΣ ΑΝΑ ΕΤΟΣ (ΛΕΞΙΚΟΛΟΓΙΚΗ ΠΡΟΣΕΓΓΙΣΗ)



Στο Διάγραμμα 4.1.Δ παρατηρείται μια θερμοχρωματική (heatmap) αναπαράσταση των ποσοστών κάθε κατηγορίας του λεξιλογίου ανά έτος. Εκτός από την κυριαρχία των αρνητικών όρων, αυτή η ανάλυση αποκαλύπτει σημαντικές μετατοπίσεις στον τρόπο έκφρασης της FOMC. Παρατηρείται ότι τα ποσοστά αβεβαιότητας είναι ιδιαίτερα υψηλά κατά τα έτη 2019 και 2024, δηλαδή σε περιόδους οικονομικής μετάβασης. Επιπλέον παρατηρείται

και η αισθητή αύξηση της τροπικής γλώσσας το 2023, δηλαδή λέξεων όπως "must", "will", το οποίο πιθανώς αντικατοπτρίζει την ανάγκη για κατηγορηματικές δηλώσεις στήριξης κατά την διάρκεια της τραπεζικής κρίσης. Αντίθετα, το 2024 παρατηρείται μια γενική μείωση τόσο των ισχυρών τροπικών όσο και νομικών όρων, υποδηλώνοντας επιστροφή σε έναν πιο προσεκτικό και αναμορφωτικό λεκτικό τρόπο.

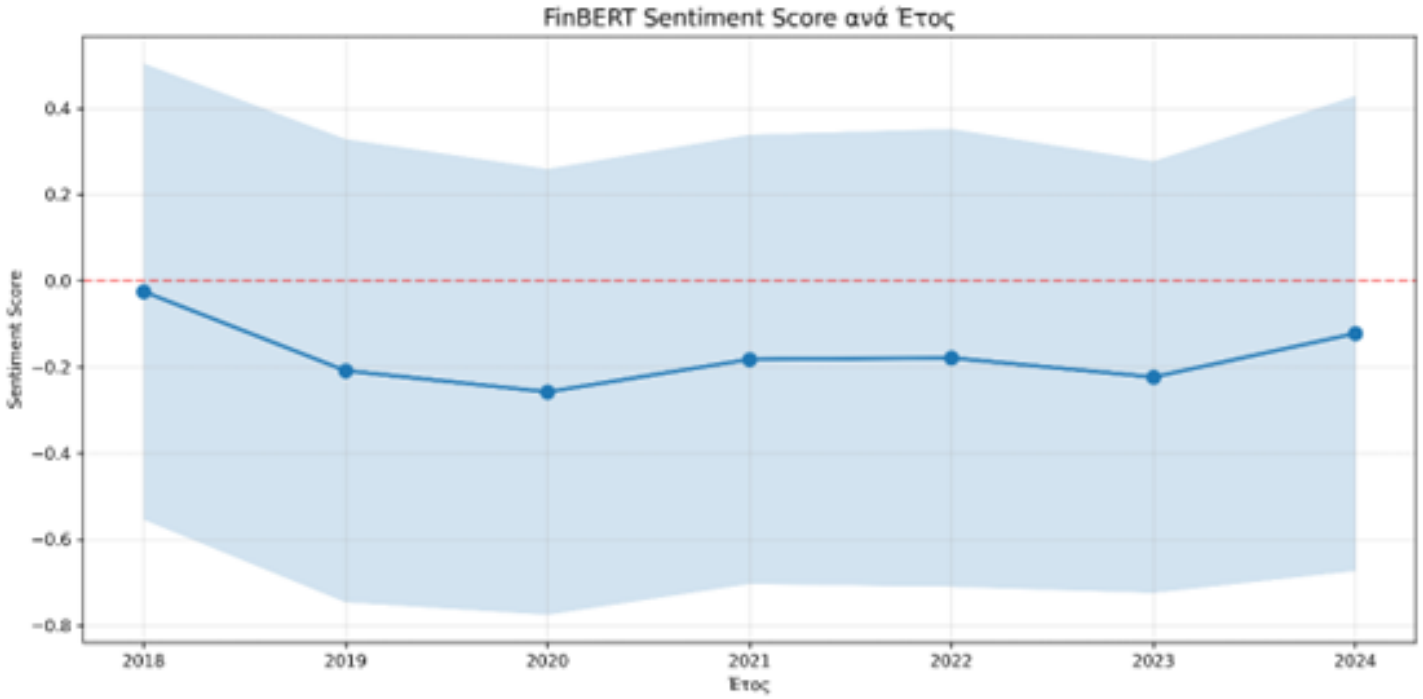
4.1.2 ΠΡΟΣΕΓΓΙΣΗ FINBERT

Η εφαρμογή του γλωσσικού μοντέλου FinBERT, ανέδειξε με μεγαλύτερη εμβάθυνση και ένταση την εικόνα του συναισθηματικού τόνου. Η ανάλυση εφαρμόστηκε σε 2378 παραγράφους και το μέσο Net Sentiment Score ήταν -0.01729, δηλαδή είναι πιο αρνητικό από αυτό του Lexicon. Αυτή η διαφορά δεν αποτελεί αντίφαση αλλά αποδεικνύει την συμπληρωματικότητα των μεθόδων. Αφού το λεξικό μετράει το στατικό φορτίο λέξεων, το FinBERT συλλαμβάνει το νόημα ολόκληρων προτάσεων, όπου ο προβληματισμός για κινδύνους και προϋποθέσεις δίνει στο επιχείρημα μια ουσιαστικά προληπτική χροιά.

ΠΙΝΑΚΑΣ 4.1.Ε: ΕΤΗΣΙΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ(FINBERT)

ΕΤΟΣ	ΜΕΣΟ SCORE	% ΘΕΤΙΚΕΣ	% ΑΡΝΗΤΙΚΕΣ	% ΟΥΔΕΤΕΡΕΣ
2018	-0,0249	27,5%	43,2%	29.3%
2019	-0,2086	17,9%	50,0%	32.1%
2020	-0,2578	12,4%	55,3%	32.3%
2021	-0,1824	19,9%	48,5%	31.6%
2022	-0,1789	20,1%	47,3%	32.6%
2023	-0,2234	15,1%	51,1%	33.8%
2024	-0,1222	23,3%	44,5%	32.2%

ΔΙΑΓΡΑΜΜΑ 4.1.ΣΤ: ΧΡΟΝΙΚΗ ΕΞΕΛΙΞΗ ΣΥΝΑΙΣΘΗΜΑΤΟΣ (FINBERT)

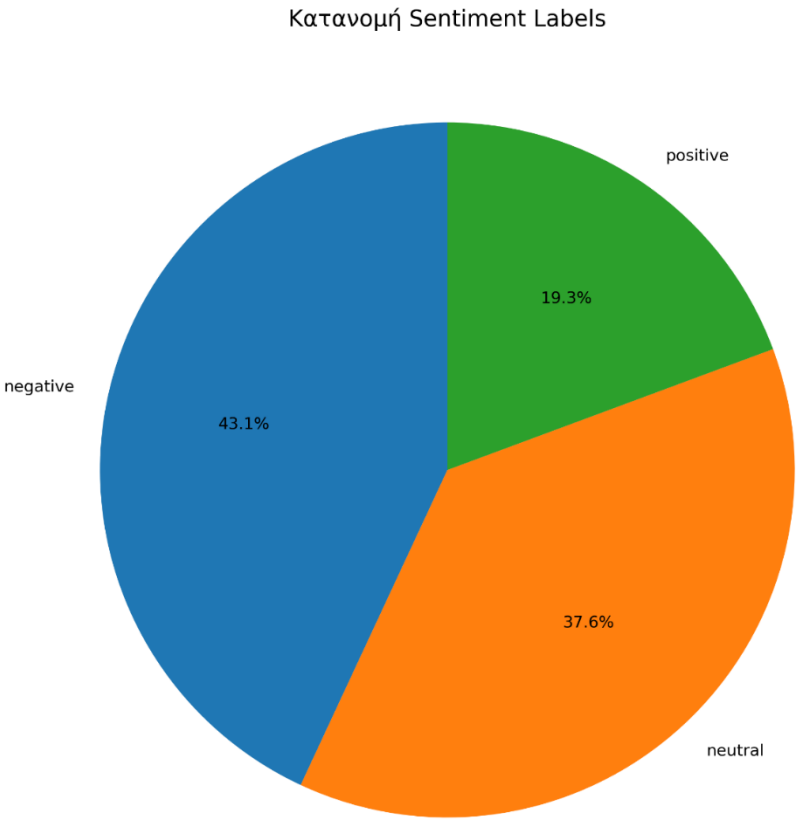


Ο Πίνακας 4.1.Ε και το Διάγραμμα 4.1.ΣΤ παρουσιάζουν την χρονική εξέλιξη του μέσου Sentiment Score σύμφωνα με το FinBERT. Τα αποτελέσματα επιβεβαιώνουν τις τάσεις που εντοπίστηκαν και με το Lexicon αλλά με μεγαλύτερο εύρος διακύμανσης και έντονη αρνητικότητα. Συγκεκριμένα:

- **2020:** Καταγράφεται το πιο αρνητικό score (-0.2578), με πάνω από 55% των παραγράφων να χαρακτηρίζονται ως αρνητικές, αντικατοπτρίζοντας το απόγειο της κρίσης και της έκτακτης αβεβαιότητας της πανδημίας.
- **2021:** Παρατηρείται σημαντική βελτίωση (score: -0.1824), με το ποσοστό θετικών λέξεων να ανεβαίνει στο 19.9%, υποδεικνύοντας την αισιοδοξία της ανάκαμψης μετά τα εμβόλια.
- **2022-2023:** Το αρνητικό συναίσθημα παραμένει έντονο (scores: -0.1789 και -0.2234), αντανακλώντας την πίεση του πληθωρισμού και τις τραπεζικές αναταραχές. Το 2023 σημειώνεται νέα άνοδος στην αρνητικότητα.

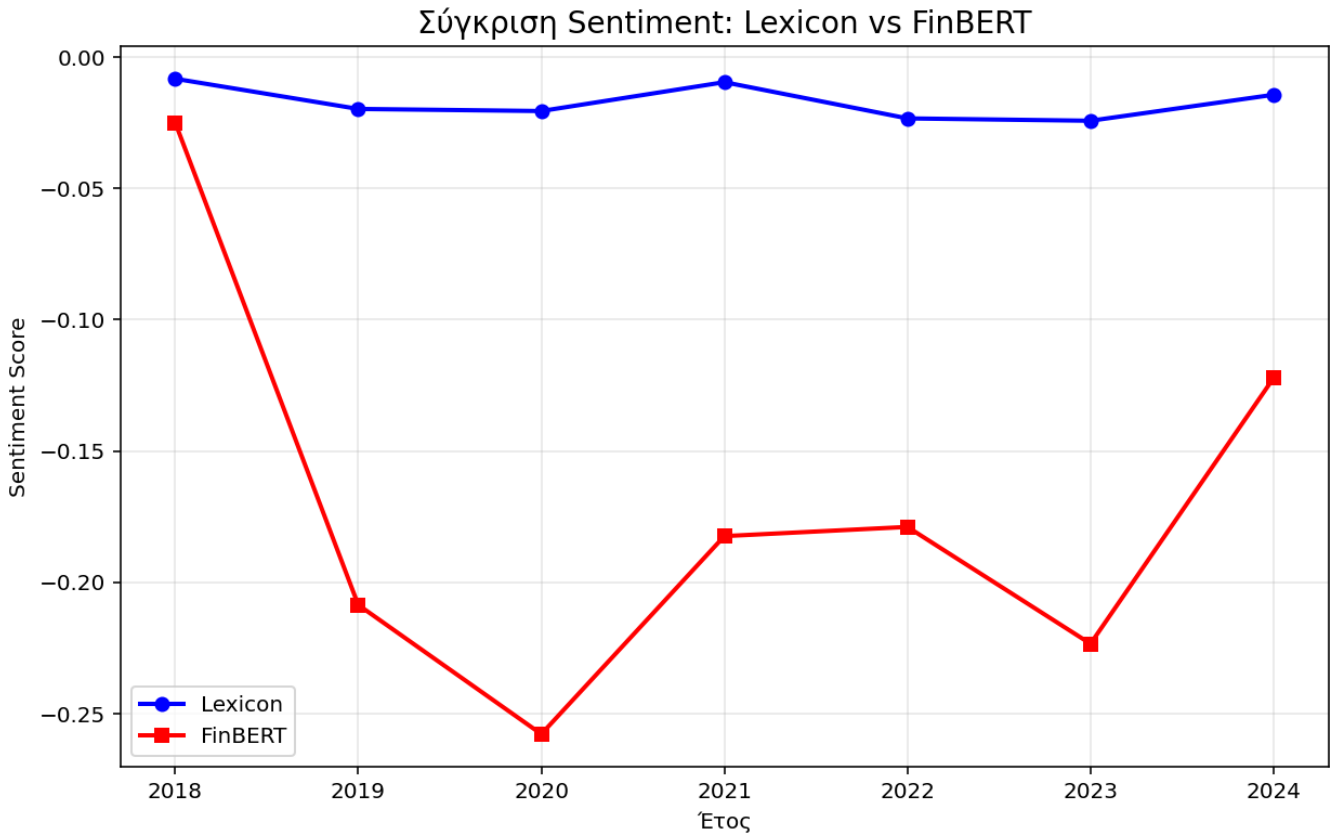
- 2024: Εμφανίζονται σημάδια σταθεροποίησης (score: -0.1222). Το ποσοστό θετικών λέξεων ανεβαίνει στο 23.3% – το υψηλότερο της περιόδου μετά το 2018 – υποδηλώνοντας επιστροφή σε πιο ισορροπημένους τόνους συζήτηση

ΣΧΗΜΑ 4.1.Ζ: ΕΤΗΣΙΑ ΣΤΑΤΙΣΤΙΚΑ ΣΥΝΑΙΣΘΗΜΑΤΟΣ(FINBERT)



Η εσωτερική δομή των πρακτικών, όπως αποκαλύπτεται από το FinBERT, είναι χαρακτηριστική (Σχήμα 4.1.Ζ). Περίπου το 43.1% των παραγράφων είναι αρνητικές (ανάλυση κινδύνων και προκλήσεων), το 37.6% ουδέτερες (περιγραφική και αναλυτική γλώσσα δεδομένων) και μόλις το 19.3% θετικές (περιγραφές ανθεκτικότητας ή θετικών εξελίξεων). Αυτή η κατανομή τονίζει ότι ο πυρήνας της εργασίας της Επιτροπής είναι η ανάλυση και διαχείριση κινδύνων, πέρα από την έκφραση ξεκάθαρου συναισθήματος.

ΔΙΑΓΡΑΜΜΑ 4.1.Η: ΣΥΓΚΡΙΣΗ FINBERT ΜΕ LEXICON (2018-2024)



Το Διάγραμμα 4.1.Η συγκρίνει τα δύο scores ανά έτος. Αποκαλύπτει ότι το FinBERT περιέχει σταθερά πιο αρνητικές τιμές. Αυτό δεν είναι αντίφαση, αλλά είναι απόδειξη της συμπληρωματικότητας των μεθόδων. Το Lexicon μετράει το φορτίο συγκεκριμένων λέξεων. Το ότι το score είναι πιο κοντά στο μηδέν υποδηλώνει μια ισορροπία στη χρήση θετικών/αρνητικών όρων μέσα στα πρακτικά. Αντίθετα το FinBERT κρίνει το νόημα ολόκληρων προτάσεων. Το πιο αρνητικό score δείχνει ότι όταν οι προτάσεις αυτές συνδυάζονται για να σχηματίσουν επιχειρήματα, το συνολικό μήνυμα τείνει προς την πρόβλεψη κινδύνων και την διαχείριση αρνητικών σεναρίων.

Συμπερασματικά, η ανάλυση με το FinBERT προσφέρει βάθος και επιβεβαιώνει ότι τα πρακτικά της FOMC όχι μόνο χρησιμοποιούν περισσότερους αρνητικούς όρους, αλλά ότι η ουσιαστική δομή των συζητήσεων είναι προσανατολισμένη στην αντιμετώπιση κινδύνων. Οι έντονες διακυμάνσεις του score ανταποκρίνονται άμεσα στις οικονομικές

κρίσεις, ενώ η υψηλή αναλογία των ουδέτερων παραγράφων υπογραμμίζει τον ρόλο της αντικειμενικής ανάλυσης.

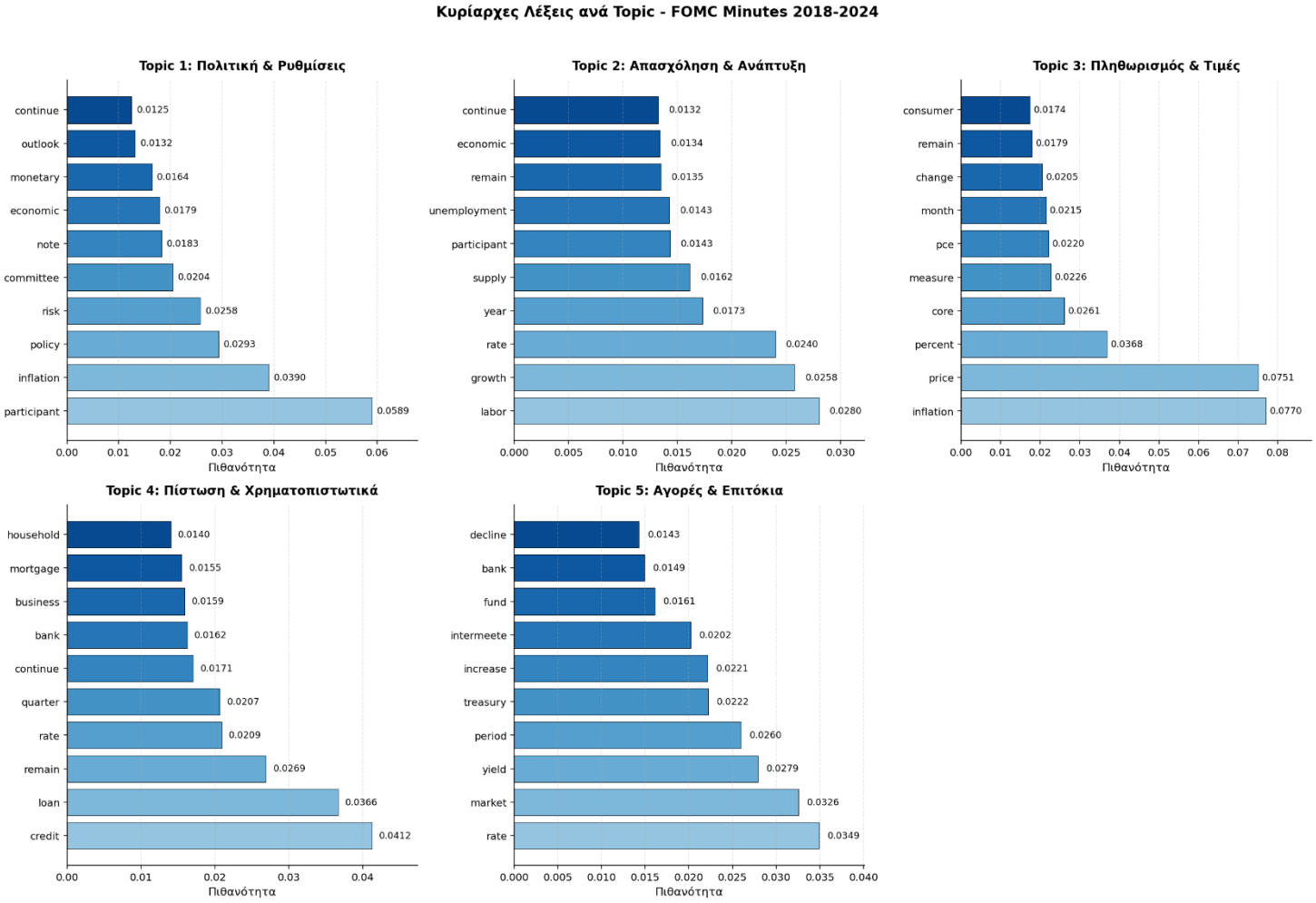
4.2 ΑΠΟΤΕΛΕΣΜΑΤΑ ΘΕΜΑΤΙΚΗΣ ΑΝΑΛΥΣΗΣ

Η θεματική ανάλυση με τη μέθοδο Latent Dirichlet Allocation (LDA) εφαρμόστηκε σε δύο επίπεδα για μια ολοκληρωμένη κατανόηση: αρχικά σε επίπεδο παραγράφου για τον εντοπισμό των σταθερών θεματικών πυλώνων, και στη συνέχεια σε επίπεδο έτους για την ανίχνευση της χρονικής εξέλιξης και της ανταπόκρισης της FOMC σε εξωτερικά σοκ.

4.2.1 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ

Η ανάλυση σε επίπεδο παραγράφου επεξεργάστηκε 1876 παραγράφους και εντόπισε πέντε διακριτά και ερμηνεύσιμα θέματα, που αντιπροσωπεύουν τους σταθερούς θεματικούς πυλώνες που διατρέχουν τα πρακτικά της ανεξάρτητα από την συγκεκριμένη χρονική περίοδο.

ΔΙΑΓΡΑΜΜΑ 4.2.Α: ΚΥΡΙΑΡΧΕΣ ΛΕΞΕΙΣ ΑΝΑ TOPIC



ΣΧΗΜΑ 4.2.Β: WORDCLOUD ΤΩΝ 5 ΘΕΜΑΤΩΝ ΤΗΣ ΑΝΑΛΥΣΗΣ LDA ΑΝΑ ΠΑΡΑΓΡΑΦΟ



Παράλληλα με τη γραφική αναπαράσταση των bar chart, στο Σχήμα 4.2.Β παρουσιάζεται ένα wordcloud που επισημαίνει οπτικά την συχνότητα και τη σημασία των λέξεων-κλειδιών για κάθε ένα από τα πέντε θέματα. Η

μεγέθυνση των λέξεων όπως “inflation”, “economic” και “policy” επιβεβαιώνει την κεντρική τους θέση στις συζητήσεις της FOMC. Από το Διάγραμμα 4.2.Α προκύπτουν οι δέκα πιο χαρακτηριστικές λέξεις για κάθε ένα από τα πέντε θέματα που προέκυψαν από την ανάλυση LDA σε επίπεδο παραγράφου.

ΘΕΜΑ 1: Επικεντρώνεται στο ευρύτερο πλαίσιο της οικονομικής πολιτικής, με κυρίαρχες λέξεις: “economic”, “monetary”, “outlook”, “committee” και “continue”. Αυτό το θεμελιώδες πλαίσιο λειτουργεί ως υπόβαθρο στο οποίο πάνω χτίζονται όλες οι άλλες αναλύσεις.

ΘΕΜΑ 2: Αφορά τις συνθήκες της αγοράς εργασίας και τις μακροοικονομικές τάσεις ανάπτυξης, με κυρίαρχες λέξεις: “labor”, “growth”, “unemployment”, “supply” και “rate”, οι οποίες δείχνουν την συστημική παρακολούθηση των δεικτών απασχόλησης και της οικονομικής δραστηριότητας.

ΘΕΜΑ 3: Πρόκειται για το πιο έντονο και συνεκτικό θέμα, που αντικατοπτρίζει την πρωταρχική ανησυχία της Επιτροπής, με κυρίαρχες λέξεις: “inflation”, “price”, “precent” , “PCE”, “core” και “consumer”, οι οποίες επισημαίνουν την εντατική παρακολούθηση των πληθωριστικών πιέσεων και των διαφόρων δεικτών τιμών.

ΘΕΜΑ 4: Αφορά την “υγεία” της πιστωτικής αγοράς και των νοικοκυριών, με κυρίαρχες λέξεις: “credit”, “loan”, “mortgage”, “household” και “bank”. Επομένως, εστιάζουν στο αν το τραπεζικό σύστημα μπορεί να παρέχει πιστώσεις και πόσο ανθεκτικό είναι.

ΘΕΜΑ 5: Σχετίζεται με τις χρηματιστηριακές αγορές και την διαχείριση των επιτοκίων, με κυρίαρχες λέξεις: “market”, “yield”, “treasury”, “rate”, “increase” και “fund”, οι οποίες εκφράζουν την παρακολούθηση των συνθηκών χρηματοδότησης και της μετάδοσης της νομισματικής πολιτικής.

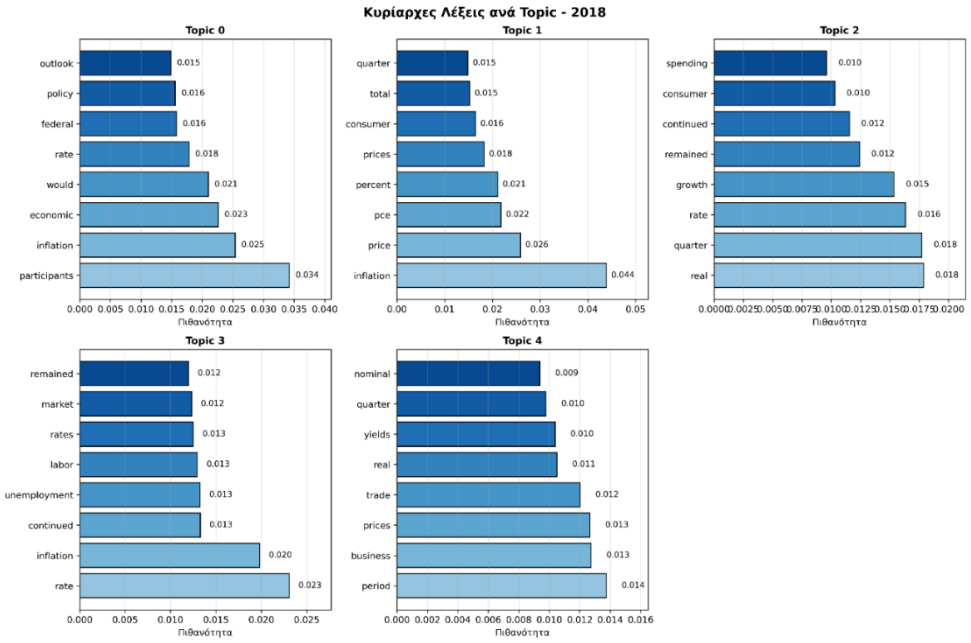
Η ανάλυση ανά παράγραφο αποκαλύπτει λοιπόν μια σταθερή και δομημένη θεματική αρχιτεκτονική στα πρακτικά της FOMC. Τα θέματα αυτά υπογραμμίζουν τις διαχρονικές προτεραιότητες της Επιτροπής.

4.2.2 ΑΝΑΛΥΣΗ LDA ΑΝΑ ΕΤΟΣ

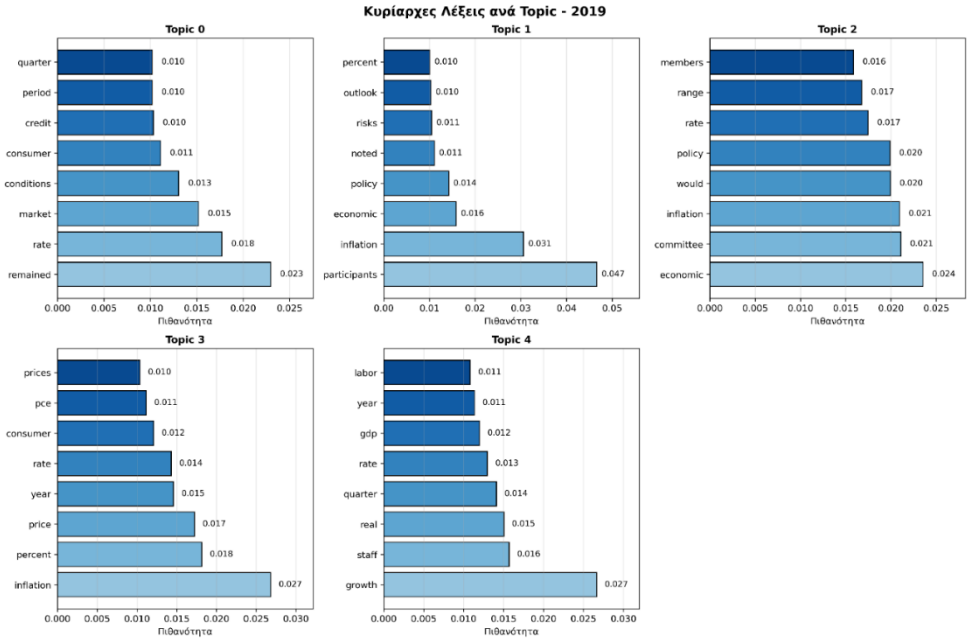
Πέρα από την σημασία των θεματικών που περιέχουν τα πρακτικά της FOMC, έχει σημασία και πως αυτές προσαρμόζονται στις ιστορικές συνθήκες. Για να μπορέσει να αναλυθεί αυτό χρειάστηκε να εφαρμοστεί LDA ξεχωριστά για κάθε έτος της περιόδου 2018-2024. Παρακάτω θα παρουσιαστούν τα ευρήματα της ανάλυσης, με τα πέντε topics και τις κυρίαρχες λέξεις ανά έτος στα Διαγράμματα 4.2.Γ με 4.2.Θ , καθώς και περιγραφή των συμπερασμάτων που εξάγονται για κάθε χρονική περίοδο ξεχωριστά.

4.2.2.1 ΠΡΟ-ΠΑΝΔΗΜΙΚΗ ΠΕΡΙΟΔΟΣ(2018-2019)

ΔΙΑΓΡΑΜΜΑ 4.2.Γ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2018



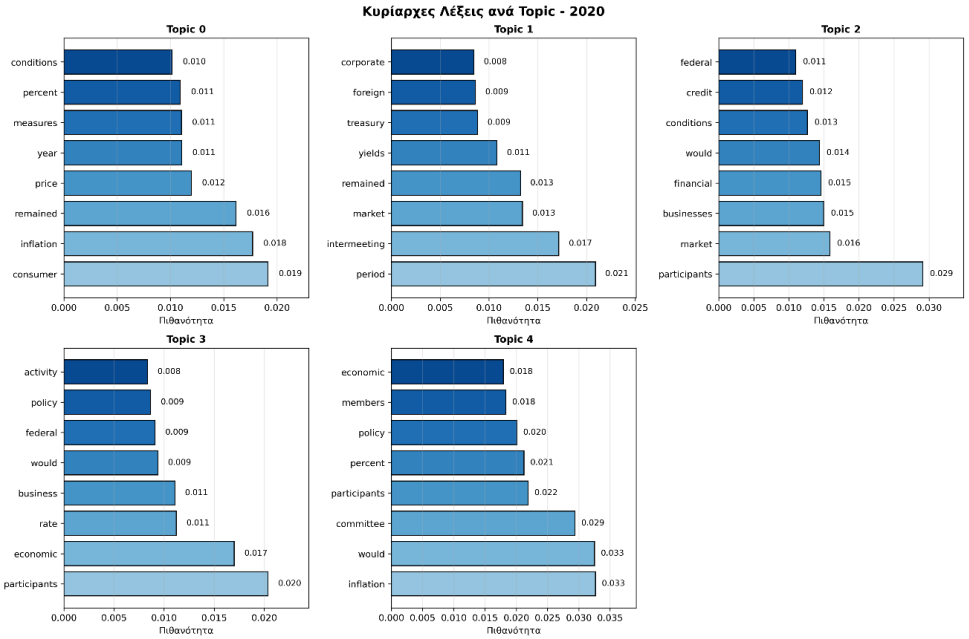
ΔΙΑΓΡΑΜΜΑ 4.2.Δ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2019



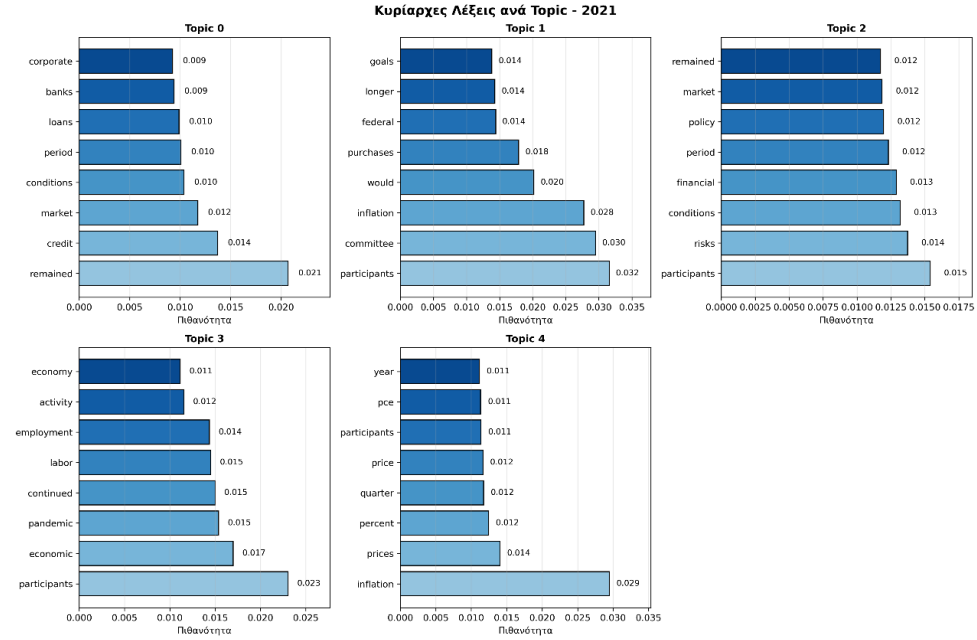
Την περίοδο πριν την εμφάνιση της πανδημίας, δηλαδή τα έτη 2018-2019, από τα Διαγράμματα 4.2.Γ και 4.2.Δ παρατηρείται ότι περίπου το 60% των θεμάτων αφορούσαν άμεσα τον πληθωρισμό, αυτό συμπεραίνεται από την εμφάνιση λέξεων όπως "inflation", "price", "PCE". Επιπλέον, το 2019 προστίθενται λέξεις όπως "growth" και "GDP", αντικατοπτρίζοντας την εστίαση σε μια ανθεκτική και αναπτυσσόμενη οικονομία πριν το σοκ.

4.2.2.2 ΠΕΡΙΟΔΟΣ ΠΑΝΔΗΜΙΑΣ COVID-19(2020-2021)

ΔΙΑΓΡΑΜΜΑ 4.2.Ε: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2020



ΔΙΑΓΡΑΜΜΑ 4.2.ΣΤ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2021

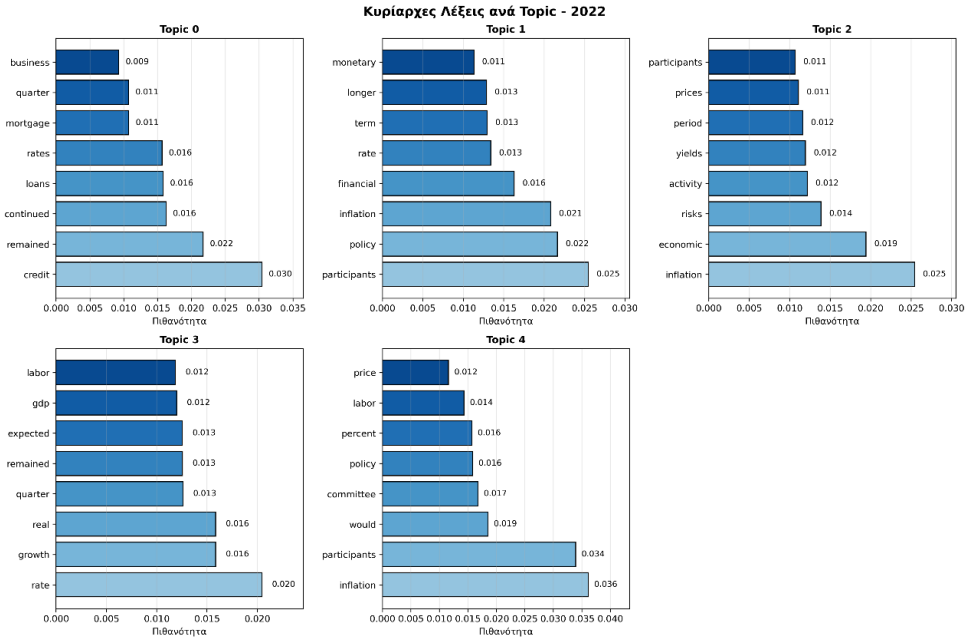


Από το Διάγραμμα 4.2.Ε, το 2020 παρατηρείται μια σημαντική μετατόπιση των προτεραιοτήτων της Επιτροπής, καθώς ξεκινάει η πανδημία COVID-19.

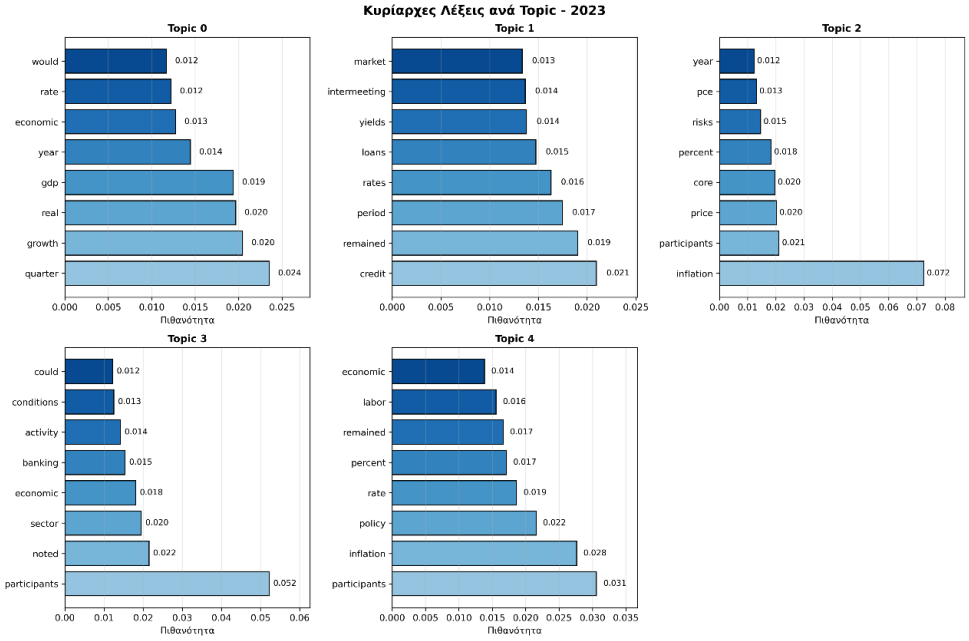
Το ποσοστό των θεμάτων που αφορούσαν τον πληθωρισμό μειώνεται στο 40%, ενώ εμφανίζονται δύο νέα έντονα θέματα που αντικατοπτρίζουν την έκτακτη κατάσταση. Το ένα εστιάζει σε χρηματοπιστωτικές συνθήκες και αγορές ("market", "yields", "intermeeting"), ενώ το άλλο περιλαμβάνει λέξεις όπως "pandemic" τονίζοντας την κεντρικότητα της κρίσης στις συζητήσεις. Το 2021 από το Διάγραμμα 4.2.ΣΤ συμπεραίνεται ότι είναι μια μεταβατική περίοδος ανάκαμψης και συγκεκριμένα η Επιτροπή αρχίζει να επανεξετάζει τις πολιτικές της. Ο πληθωρισμός παραμένει ένα από τα κύρια θέματα περίπου σε ποσοστό 40%, αλλά εμφανίζεται ένα ξεχωριστό θέμα που συζητά την ποσοτική χαλάρωση ("purchases"), ενώ η λέξη "pandemic" εξακολουθεί να εμφανίζεται αλλά σε μειωμένο βαθμό.

4.2.2.3 ΜΕΤΑ-ΠΑΝΔΗΜΙΚΗ ΚΑΙ ΠΛΗΘΩΡΙΣΤΙΚΗ ΠΕΡΙΟΔΟΣ(2022-2023)

ΔΙΑΓΡΑΜΜΑ 4.2.Ζ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2022



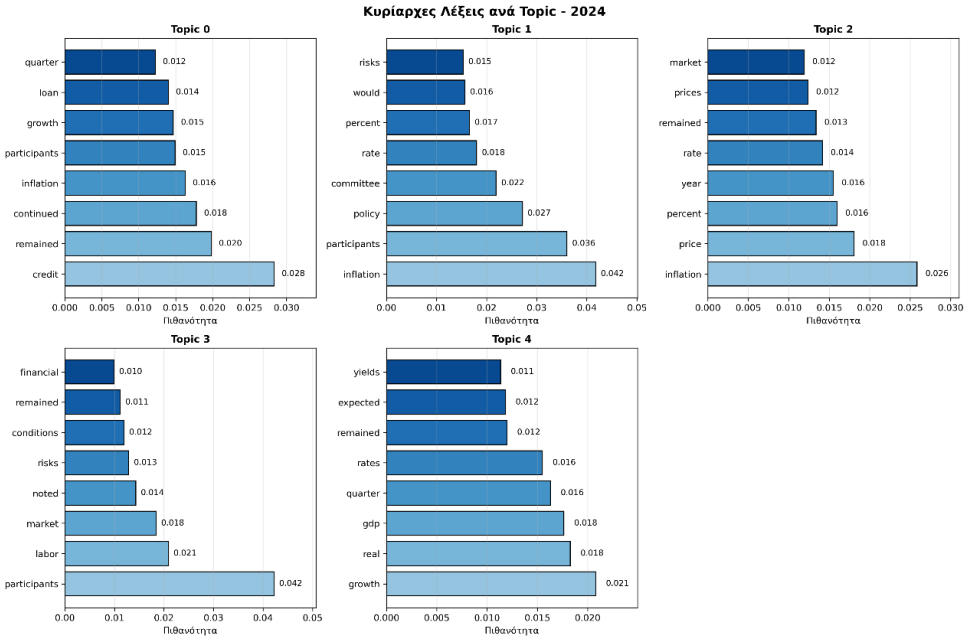
ΔΙΑΓΡΑΜΜΑ 4.2.Η: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2023



Το 2022-2023 είναι έτη που σηματοδοτούν την επιστροφή και ενίσχυση της εστίασης στο πληθωρισμό. Συγκεκριμένα, το 2022(Διάγραμμα 4.2.Ζ) το 60% των θεμάτων αφορά άμεσα τον πληθωρισμό και την νομισματική πολιτική ("monetary", "mortgage"), αντανακλώντας την έναρξη του κύκλου αυστηροποίησης. Το 2023 (Διάγραμμα 4.2.Η) ενώ ο πληθωρισμός παραμένει κύριο θέμα σε ποσοστό 40%, εμφανίζεται ένα νέο διακριτό θέμα που αναφέρεται ρητά στον τραπεζικό τομέα "banking", "sector", "core" πιθανώς σε ανταπόκριση της άνοιξης εκείνου του έτους.

4.2.2.4 ΠΕΡΙΟΔΟΣ ΣΤΑΘΕΡΟΠΟΙΗΣΗΣ(2024)

ΔΙΑΓΡΑΜΜΑ 4.2.Θ: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ – 2024



Τέλος, από το Διάγραμμα 4.2.Θ συμπεραίνουμε ότι το 2024 θα μπορούσε να χαρακτηριστεί ως μια περίοδο σταθεροποίησης. Ο πληθωρισμός παραμένει σημαντικός. Ταυτόχρονα, ένα νέο θέμα κερδίζει έμφαση το οποίο ήταν η αγορά εργασίας ("labor", "market"). Αυτή η μετατόπιση δείχνει ότι, καθώς οι πληθωριστικές πιέσεις αρχίζουν να υποχωρούν, η προσοχή της Επιτροπής επιστρέφει στην παρακολούθηση της αγοράς εργασίας.

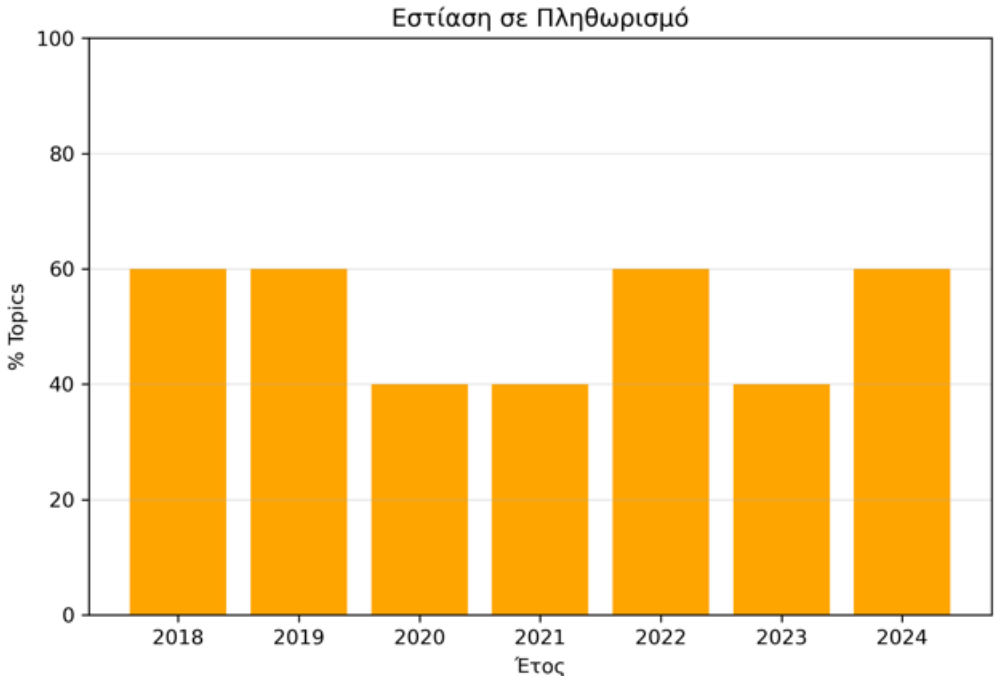
ΠΙΝΑΚΑΣ 4.2.Η: ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ ΑΝΑ ΕΤΟΣ

ΕΤΟΣ	ΚΥΡΙΑΡΧΑ ΘΕΜΑΤΑ	%TOPICS ΓΙΑ ΠΛΗΘΩΡΙΣΜΟ	ΒΑΣΙΚΕΣ ΛΕΞΕΙΣ
2018	Πληθωρισμός, Πολιτική	60%	inflation, price, PCE
2019	Πληθωρισμός, Ανάπτυξη	60%	inflation, growth, GDP
2020	Πληθωρισμός, Χρηματοπιστωτικά	40%	Market, yields, pandemic

2021	Πληθωρισμός, Πανδημία	40%	inflation, purchases, pandemic
2022	Πληθωρισμός, Χρηματοπιστωτικά	60%	inflation, monetary, mortgage
2023	Πληθωρισμός, Τραπεζικά	40%	banking, sector, core
2024	Πληθωρισμός, Εργασία	60%	inflation, labor, market

Από αυτή την ανάλυση αυτή προκύπτει ότι ένα από τα θέματα που συζητούνται ανεξαρτήτως του έτους είναι ο πληθωρισμός. Αυτό που παρατηρείται όμως είναι ότι στα έτη που επικρατούν κρίσεις μειώνεται το ποσοστό εμφάνισης τους. Στο Διάγραμμα 4.2.Η που ακολουθεί παρατηρείται ότι τα έτη 2020-2021 που είναι η πανδημία COVID-19 και το 2023 που υπήρχαν αναταραχές στο τραπεζικό σύστημα μειώνεται το ποσοστό συζήτησης του πληθωρισμού.

ΔΙΑΓΡΑΜΜΑ 4.2.Ι: ΠΟΣΟΣΤΟ ΕΜΦΑΝΙΣΗΣ ΠΛΗΘΩΡΙΣΜΟΥ ΣΤΑ ΘΕΜΑΤΑ ΑΝΑ ΧΡΟΝΙΑ



4.3 ΣΥΝΟΛΙΚΗ ΕΡΜΗΝΕΙΑ ΑΠΟΤΕΛΕΣΜΑΤΩΝ

Τα ευρήματα τόσο της ανάλυσης συναισθήματος όσο και της θεματικής ανάλυσης δημιουργούν μια ισχυρή και πολυδιάστατη αφήγηση για την επικοινωνία και τις προτεραιότητες της FOMC κατά την περίοδο μεταξύ 2018 με 2024. Η μελέτη επιβεβαιώνει τα εξής:

Η γλώσσα των πρακτικών λειτουργεί ως ευαίσθητος δείκτης οικονομικής αβεβαιότητας και πολιτικής προσαρμογής. Η συνεχής αρνητική ή τουλάχιστον προσεκτική τροπικότητα που καταγράφηκε τόσο από το Lexicon όσο και από το FinBERT δεν αποτελεί τυχαία γλωσσική επιλογή. Αντίθετα, ανταποκρίνεται στην ενδογενή λειτουργία της Επιτροπής ως οργάνου διαχείρισης κινδύνων. Το FinBERT με την ικανότητα του να κατανοεί συμφραζόμενα αποκαλύπτει ότι αυτή η αρνητικότητα δεν είναι απλώς υπερβολική χρήση αρνητικών λέξεων, αλλά θεμελιώδες δομικό χαρακτηριστικό των συζητήσεων, όπου η ανάλυση των κινδύνων και των προκλήσεων βρίσκεται στο επίκεντρο. Παράλληλα η θεματική ανάλυση LDA χαρτογραφεί τι συζητείται. Η διαχρονική σταθερότητα των θεμάτων όπως πληθωρισμός, απασχόληση δείχνει πυρήνες συνεχούς προσοχής, ενώ μετατοπίσεις στο βάρος τους (πχ. άνοδος των χρηματοπιστωτικών θεμάτων το 2023, άνοδος εργασίας το 2024) αποτελούν δείκτες της προσαρμοστικότητας και της ανταπόκρισης της Επιτροπής σε εξωγενείς σοκ.

Η παρούσα εργασία καταδεικνύει πως ο συναισθηματικός τόνος των πρακτικών είναι αρνητικός/προσεκτικός και μεταβάλλεται δυναμικά, φτάνοντας σε χαμηλά κατά την διάρκεια της πανδημίας COVID-19 και της πληθωριστικής κρίσης του 2022-2023. Επίσης, οι θεματικές προτεραιότητες εξελίσσονται σαφώς από την προ-πανδημική εστίαση στον πληθωρισμό και την ανάπτυξη, προς την διαχείριση της πανδημίας και των χρηματοπιστωτικών αγορών (2020-2021) και στη συνέχεια στην επανεστίαση στον πληθωρισμό και στη διαχείριση των τραπεζικών αναταραχών (2022-2023), πριν επιστρέψει σε μια πιο ισορροπημένη εστίαση που περιλαμβάνει και την αγορά εργασίας (2024). Επομένως, η χρονική εξέλιξη τόσο του συναισθήματος όσο και των θεμάτων συσχετίζεται άμεσα με σημαντικά ιστορικά και οικονομικά γεγονότα, όπως η πανδημία COVID-19 και η πληθωριστική κρίση. Αυτό υποστηρίζει ότι τα πρακτικά δεν είναι μόνο τεχνικά έγγραφα, αλλά και επικοινωνιακά εργαλεία που αντικατοπτρίζουν και διαμορφώνουν την οικονομική αφήγηση.

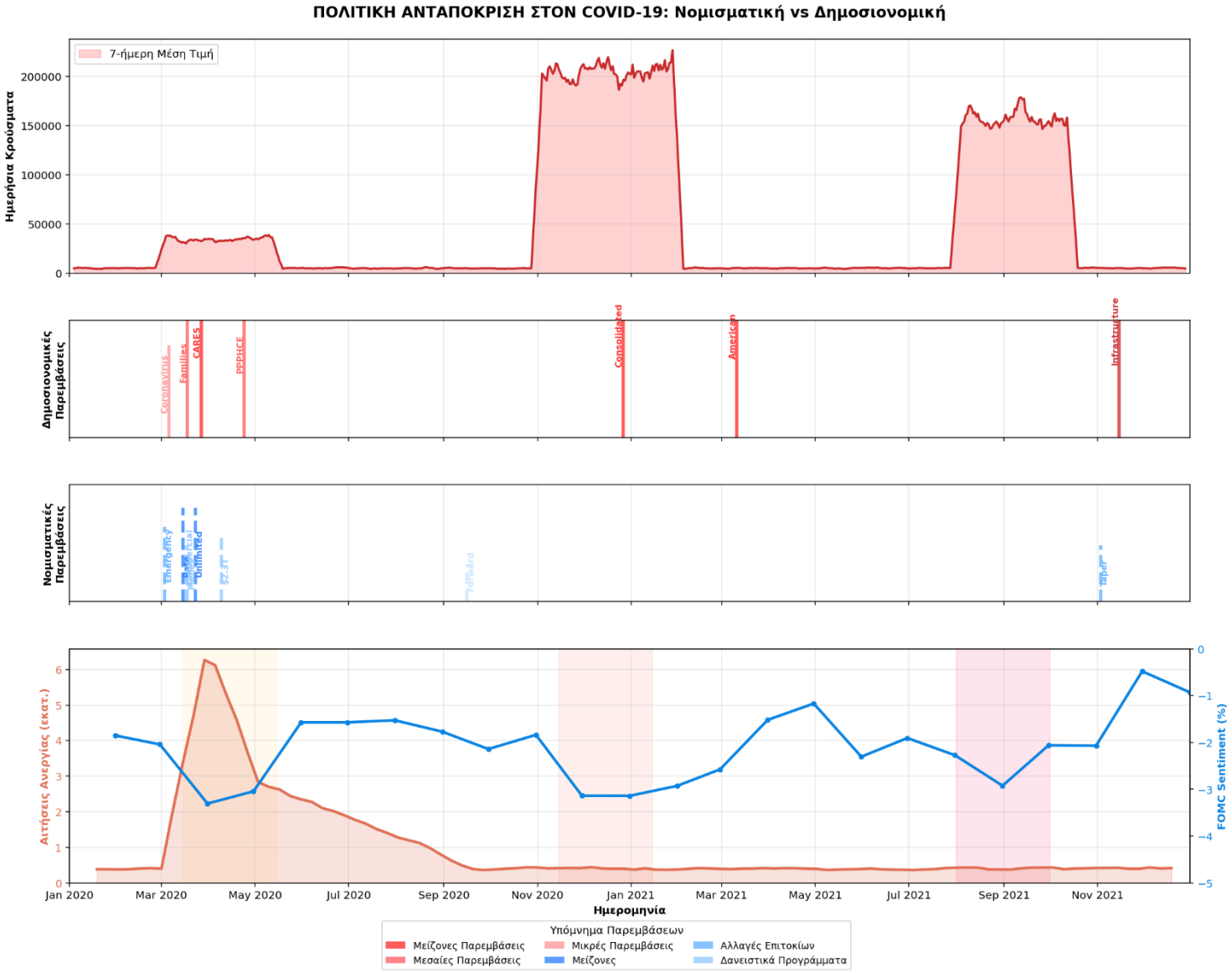
4.4 ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ: ΣΥΝΔΥΑΣΜΟΣ ΓΛΩΣΣΙΚΩΝ ΕΥΡΗΜΑΤΩΝ ΚΑΙ ΔΕΙΚΤΩΝ ΠΟΛΙΤΙΚΗΣ

Η ανάλυση που προηγήθηκε έδειξε πώς ο συναισθηματικός τόνος και οι θεματικές προτεραιότητες των FOMC Minutes μεταβάλλονται δυναμικά. Η παρόν υποενότητα στοχεύει να τοποθετήσει αυτά τα γλωσσικά ευρήματα στο χρονικό και πολιτικό πλαίσιο στο οποίο παραχθήκαν. Η ανάλυση επικεντρώνεται στην περίοδο έντονης κρίσης (2020-2021) για να διερευνηθεί ο χρονικός συντονισμός μεταξύ της γλωσσικής έκφρασης και της πραγματικής πολιτικής δράσης , αλλά και το κατά πόσο η γλώσσα των πρακτικών μπορεί να λειτουργήσει ως δείκτης του μεγέθους ή της αποτελεσματικότητας της πολιτικής ανταπόκρισης.

4.4.1 ΧΡΟΝΟΛΟΓΙΟ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ ΣΤΟΝ COVID-19

Με βάση την ανάλυση δεδομένων των πολιτικών παρεμβάσεων προκύπτει το Διάγραμμα 4.4.Α που ακολουθεί. Στο οποίο μπορούν να διακριθούν τρεις φάσεις ανταπόκρισης στον COVID-19.

Διάγραμμα 4.4.Α: ΧΡΟΝΟΛΟΓΙΟ ΠΟΛΙΤΙΚΗΣ ΑΝΤΑΠΟΚΡΙΣΗΣ ΣΤΟΝ COVID-19(2020-2021) - ΝΟΜΙΣΜΑΤΙΚΕΣ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΕΣ ΠΑΡΕΜΒΑΣΕΙΣ



Οι φάσεις που παρατηρούνται είναι οι εξής:

Φάση 1: Επείγουσα Ανταπόκριση (Μάρτιος-Απρίλιος 2020)

Χαρακτηρίζεται από ταυτόχρονες, εκτάκτου μεγέθους παρεμβάσεις και από τις δύο πλευρές. Η FOMC στις συνεδριάσεις της 15 και 23 Μαρτίου, μείωσε τα επιτόκια στο 0-0,25% και ανακοίνωσε "απεριόριστη" ποσοτική χαλάρωση. Αυτές οι αποφάσεις καταγράφηκαν αμέσως στα πρακτικά της εποχής με έναν τόνο έκτακτης ανάγκης, με την χρήση λέξεων όπως "unprecedented" και "powerful support". Σε ελάχιστη χρονική απόσταση

και συγκεκριμένα στις 27 Μαρτίου, η πολιτική ηγεσία υπέγραψε τον CARES Act (€2.2 τρισ.),η μεγαλύτερη δημοσιονομική παρέμβαση στην αμερικάνικη ιστορία. Αυτός ο χρονικός συντονισμός υποδηλώνει μια σκόπιμα συλλογική προσέγγιση "όλων των κατευθύνσεων" για την αποφυγή οικονομικής κατάρρευσης.

Φάση 2: Στήριξη και Σταθεροποίηση (Μάιος-Δεκέμβριος 2020)

Κατά τη διάρκεια αυτής της περιόδου και οι δύο πολιτικές συνεχίστηκαν με συμπληρωματικά μέτρα (πχ. Paycheck Protection Program, Main Street Lending Program). Το συναίσθημα στα πρακτικά αν και παρέμεινε βαθιά αρνητικό, σταθεροποιήθηκε σε ελαφρώς υψηλότερα επίπεδα (από -3.31% σε -2.5% περίπου), υποδηλώνοντας ότι η φάση των έντονων έκτακτων μέτρων είχε υποχωρήσει.

Φάση 3: Προετοιμασία για Μεταβίβαση (2021)

Με την άφιξη των εμβολίων και την έναρξη ανάκαμψης, το πολιτικό προσκήνιο άλλαξε. Η ψήφιση του American Rescue Plan (Μάρτιος 2021) σήμανε την τελευταία μεγάλη δημοσιονομική ώθηση. Παράλληλα, οι συζητήσεις της FOMC άρχισαν να μετατοπίζονται προσεκτικά, με τα πρακτικά να εισάγουν θέματα, όπως η μείωση (tapering) των αγοράς ομολόγων. Αυτό αντανακλάται και στη βελτιωμένη, αν και ακόμα προσεκτική, συναισθηματική βαθμολογία για το 2021 (-1.82% σύμφωνα με το FinBERT).

4.4.2 ΣΥΓΚΡΙΤΙΚΗ ΑΝΑΛΥΣΗ ΠΟΛΙΤΙΚΩΝ ΠΡΟΣΕΓΓΙΣΕΩΝ

Η προηγούμενη ανάλυση έδειξε τον χρονικό συντονισμό μεταξύ νομισματικής και δημοσιονομικής πολιτικής. Για να κατανοηθεί πλήρως η σχέση μεταξύ της γλωσσικής έκφρασης των πρακτικών και της συνολικής πολιτικής ανταπόκρισης, είναι απαραίτητο να εξεταστούν οι θεμελιώδεις διαφορές στη φύση και τη λειτουργία αυτών των δύο πολιτικών τύπων, οι οποίες εξηγούνται στον **Πίνακα 4.4.Β.**

Πίνακας 4.1.Β: ΣΥΓΚΡΙΤΙΚΑ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΝΟΜΙΣΜΑΤΙΚΩΝ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΩΝ ΠΑΡΕΜΒΑΣΕΩΝ (2020-2021)

ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ	ΝΟΜΙΣΜΑΤΙΚΗ ΠΟΛΙΤΙΚΗ	ΔΗΜΟΣΙΟΝΟΜΙΚΗ ΠΟΛΙΤΙΚΗ
Κύριος Φορέας	Επιτροπή Ανοιχτής Αγοράς (FOMC)/ Federal Reserve	Ομοσπονδιακή Κυβέρνηση και Κογκρέσο των ΗΠΑ
Φύση Παρεμβάσεων	Συνεχείς-Διαβαθμισμένες: Συνδυασμός εργαλείων (επιτόκια, QE, δανειστικά προγράμματα)	Διακριτές-Μεγάλης Κλίμακας: Ορισμένες, μεγάλες νομοθετικές πράξεις (CARES Act)
Χρονική Δυναμική	Ταχεία Ανταπόκριση: Αποφάσεις σε ημέρες	Απαιτεί Περισσότερο Χρόνο: Νομοθετική διαδικασία σε εβδομάδες/μήνες
Αντανάκλαση στα FOMC Minutes	Άμεση και Κεντρική: Τα μέτρα είναι αντικείμενο συζήτησης και απόφασης στα ίδια τα πρακτικά	Έμμεση: Αποτελούν το πλαίσιο (background) της οικονομικής κατάστασης που αναλύει η FOMC
Κύριο Εργαλείο Ανάλυσης	Γλωσσική Ανάλυση (Sentiment, Topics): Άμεση συσχέτιση με τον τόνο και τα θέματα των συζητήσεων	Ποσοτική Ανάλυση (Χρονολόγιο, Ποσά): Ανάλυση του μεγέθους και του χρονισμού των παρεμβάσεων.

Δημοσιονομική Πολιτική:

Η δημοσιονομική απάντηση χαρακτηρίστηκε από λίγες αλλά τεράστιες σε ποσό παρεμβάσεις (CARES Act, American Rescue Plan, Consolidated Appropriations) αντιπροσωπεύουν το 74% του συνόλου των δαπανών για την περίοδο. Αυτό το χαρακτηριστικό δεν μπορεί να συλληφθεί άμεσα μέσω της ανάλυσης συναισθήματος στα FOMC Minutes. Οι συζητήσεις της Επιτροπής δεν επικεντρώνονται στη συζήτηση νομοσχέδιων, αλλά στην ανάλυση των μακροοικονομικών συνθηκών που αυτά δημιουργούν. Η ανάλυσή τους στα πρακτικά γίνεται έμμεσα, μέσω της συζήτησης για επιδράσεις σε δεδομένα όπως η απασχόληση και η κατανάλωση.

Νομισματική Πολιτική:

Αντιθέτως, η νομισματική πολιτική της FOMC υλοποιήθηκε μέσω ενός συνδυασμού εργαλείων (επιτόκια, ποσοτική χαλάρωση, δανειστικά προγράμματα) και ήταν εντός του άμεσου πεδίου των συζητήσεων που καταγράφονται στα πρακτικά. Αυτό εξηγεί γιατί τα γλωσσικά ευρήματα της παρούσας έρευνας συνδέονται άμεσα και κυρίως με αυτήν. Το μέσο συναίσθημα του 2020 (-2.59%), η χρήση λέξεων έκτακτης ανάγκης ("unprecedented", "powerful support") και η έντονη εμφάνιση θεμάτων που αφορούν τις χρηματοπιστωτικές αγορές και τη ρευστότητα αποτελούν την άμεση γλωσσική αντανάκλαση της ίδιας της διαδικασίας λήψης αυτών των αποφάσεων.

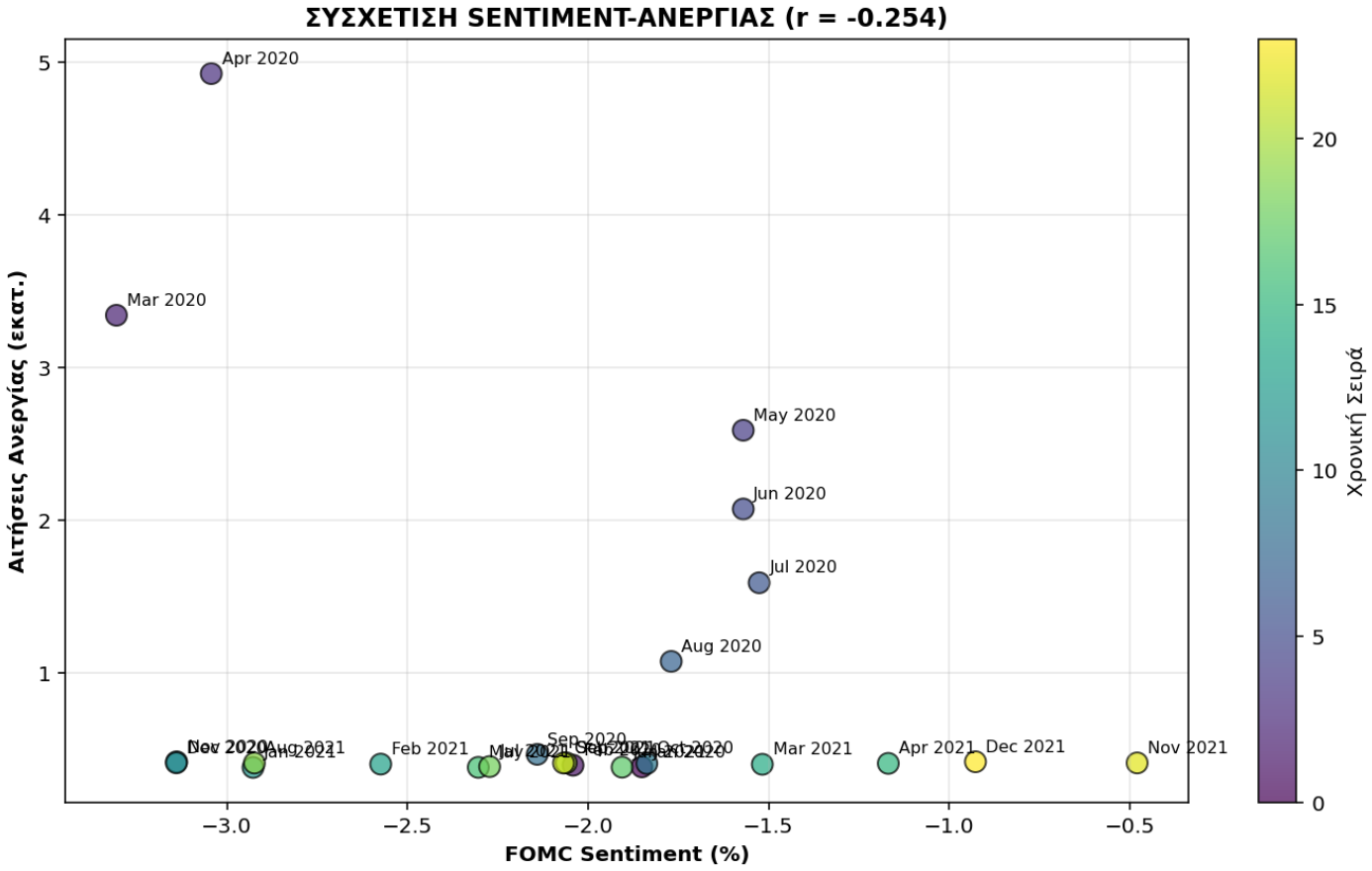
Συμπερασματικά, η ανάλυση των FOMC Minutes μέσω NLP αποτελεί ένα ισχυρό εργαλείο για την κατανόηση της νομισματικής διαδικασίας λήψης αποφάσεων και του πλαισίου σκέψης της Επιτροπής. Δεν αποτελεί ωστόσο, έναν πλήρη δείκτη της συνολικής πολιτικής ανταπόκρισης, η οποία περιλαμβάνει και την κρίσιμη, αλλά γλωσσικά πιο απομακρυσμένη, δημοσιονομική διάσταση.

4.4.3 ΓΛΩΣΣΙΚΟΣ ΤΟΝΟΣ ΩΣ ΔΕΙΚΤΗΣ ΑΒΕΒΑΙΟΤΗΤΑΣ

Η προηγούμενη ανάλυση έδειξε πώς ο γλωσσικός τόνος των FOMC Minutes μεταβάλλεται σε συγχρονισμό με τις πολιτικές φάσεις. Όμως, ένα βασικό ερώτημα παραμένει: «Μπορεί ο συναισθηματικός τόνος να λειτουργήσει όχι μόνο ως ιστορικό αρχείο, αλλά και ως προγνωστικό εργαλείο για την οικονομική πορεία;». Για να απαντηθεί αυτό, διεξήχθη στατιστική ανάλυση συσχέτισης μεταξύ του μέσου μηνιαίου συναισθήματος των πρακτικών (σύμφωνα με το FinBERT) και ενός κρίσιμου δείκτη της οικονομικής δραστηριότητας: των εβδομαδιαίων αρχικών αιτήσεων ανεργίας (initial jobless claims) για την περίοδο 2020-2021. Σκοπός αυτής της εμπειρικής δοκιμής είναι:

a) Να ελεγχθεί αν ο υποκειμενικός γλωσσικός τόνος συσχετίζεται σημαντικά με αντικειμενικά οικονομικά γεγονότα και
b) Να αξιολογηθεί εάν μια τέτοια συσχέτιση θα μπορούσε να του προσδώσει προγνωστική ισχύ.

Διάγραμμα 4.4.Γ: ΣΥΣΧΕΤΙΣΗ ΜΕΣΟΥ ΣΥΝΑΙΣΘΗΜΑΤΟΣ FOMC ΚΑΙ ΕΒΔΟΜΑΔΙΑΙΩΝ ΑΙΤΗΣΕΩΝ ΑΝΕΡΓΙΑΣ (2020-2021)



Το Διάγραμμα 4.4.Γ παρουσιάζει τη σχέση μεταξύ των δύο μεταβλητών. Ο υπολογισμός του συντελεστή συσχέτισης Pearson(r) αποκάλυψε μια ασθενή αρνητική συσχέτιση (r = -0.254). Αυτή η τιμή αν και υποδεικνύει γενική τάση, βρίσκεται πολύ μακριά από το 1(τέλεια θετική συσχέτιση) ή το -1 (τέλεια αρνητική συσχέτιση). Αυτό το στατιστικό εύρημα έχει σημαντικές προεκτάσεις για την ερμηνεία των γλωσσικών δεδομένων:

1. Κατεύθυνση Σχέσης (Αρνητικό πρόσημο):

Η αρνητική τιμή σημαίνει ότι υπάρχει μια γενική τάση όπου βελτιώσεις (λιγότερο αρνητικό) στο συναίσθημα των πρακτικών τείνουν να συσχετίζονται με πτώσεις στις αιτήσεις ανεργίας. Αυτό είναι διαισθητικά λογικό: καθώς όσο οι οικονομικές προοπτικές βελτιώνονται, τόσο η γλώσσα της FOMC γίνεται λιγότερο απαισιόδοξη όσο και οι συνθήκες στην αγορά εργασίας.

2. Ισχύς Σχέσης (Χαμηλή απόλυτη τιμή):

Η τιμή |0.254| υποδηλώνει μια πολύ ασθενή γραμμική σχέση. Στην πράξη, αυτό σημαίνει ότι η γνωρίζοντας μόνο το συναίσθημα των πρακτικών, δεν μπορούμε να κάνουμε αξιόπιστες προβλέψεις για τις μελλοντικές αιτήσεις ανεργίας, και αντίστροφα. Η συσχέτιση είναι τόσο ασθενής ώστε να μην έχει πρακτική προβλεπτική αξία.

Επομένως, το κρίσιμο συμπέρασμα που προκύπτει είναι ότι ο συναισθηματικός τόνος λειτουργεί ως ένας ανακλαστικός (coincident) και όχι προγνωστικός (leading) δείκτης. Αυτό σημαίνει ότι η γλώσσα των πρακτικών αντιπροσωπεύει μια έγκαιρη απάντηση στην υφιστάμενη οικονομική πραγματικότητα, αλλά δεν μπορεί να χρησιμοποιηθεί αξιόπιστα για την πρόβλεψη μελλοντικών οικονομικών εξελίξεων.

Απόδειξη της Ανακλαστικής Φύσης: Η σχέση μεταξύ γλώσσας και πραγματικότητας εμφανίζεται με σαφήνεια στα σημεία κρίσης. Το συναίσθημα έφτασε στο αποκορύφωμα της αρνητικότητάς του την ίδια περίοδο (Μάρτιος 2020) που η οικονομία βρισκόταν στο επίκεντρο του σοκ: οι εβδομαδιαίες αιτήσεις ανεργίας έφτασαν στις 6,5 εκατομμύριες, ένα ιστορικό ρεκόρ. Αντίστοιχα, η σταδιακή βελτίωση του συναισθήματος κατά το 2021 συνέπεσε χρονικά με την προσέλκυση των δεικτών της αγοράς εργασίας. Αυτή η άμεση χρονική σύμπτωση επιβεβαιώνει ότι οι συζητήσεις της FOMC «καθρεφτίζουν» με αξιοσημείωτη ευαισθησία την επικρατούσα οικονομική κατάσταση.

Όρια στην Προγνωστική Ικανότητα: Η ασθενής συσχέτιση (r = −0,254) αποκαλύπτει το θεμελιώδες όριο του γλωσσικού δείκτη. Η πορεία της οικονομίας και η τελική αποτελεσματικότητα των πολιτικών αποφάσεων καθορίζονται από ένα πολύπλοκο δίχτυ παραγόντων, το οποίο αποτελείται από δημογραφικά στοιχεία και παγκόσμιες αλυσίδες εφοδιασμού έως ψυχολογικά εφέ και τεχνολογικές καινοτομίες, που δεν συλλαμβάνονται πλήρως στη ρητορική μιας συνεδρίασης. Η FOMC αν και επηρεάζεται από το συναίσθημα της στιγμής, λειτουργεί μέσα σε ένα πλαίσιο πολύπλοκων οικονομετρικών μοντέλων και πολυδιάστατων δεδομένων. Η γλώσσα των πρακτικών επομένως μπορεί να αποκαλύψει το συναισθηματικό και γνωστικό πλαίσιο μέσα στο οποίο λαμβάνονται οι αποφάσεις, αλλά δεν αποτελεί από μόνη της ένα μηχανισμό πρόβλεψης για τις ίδιες τις αποφάσεις ή τις οικονομικές τους συνέπειες.

Συμπερασματικά, η ανάλυση επιβεβαιώνει την αξία της γλωσσικής προσέγγισης. Αποτελεί ένα ισχυρό διαγνωστικό εργαλείο για τη νομισματική σκέψη σε πραγματικό χρόνο, ενώ ταυτόχρονα καθορίζει με σαφήνεια τα όριά της. Τα FOMC Minutes πρέπει να αντιμετωπίζονται ως

μια μοναδική πηγή εσωτερικής προοπτικής και διαφάνειας, που μας επιτρέπει να κατανοήσουμε τα κίνητρα, τις ανησυχίες και τη λογική της Επιτροπής. Δεν πρέπει ωστόσο, να θεωρούνται ως μια αυτόνομη προγνωστική βάση δεδομένων. Η πραγματική συνεισφορά της ανάλυσης κειμένου έγκειται στην ικανότητά της να φωτίζει το «γιατί» και το «πώς» πίσω από τις αποφάσεις, προσφέροντας ένα βάθος κατανόησης που τα ποσοτικά δεδομένα μόνα τους συχνά αδυνατούν να αποκαλύψουν.

4.4.4 ΣΥΜΠΕΡΑΣΜΑΤΑ ΣΥΓΚΡΙΤΙΚΗΣ ΑΝΑΛΥΣΗΣ ΝΟΜΙΣΜΑΤΙΚΗΣ ΚΑΙ ΔΗΜΟΣΙΟΝΟΜΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

Η συγκριτική ανάλυση της νομισματικής και δημοσιονομικής πολιτικής κατά τη διάρκεια της πανδημίας COVID-19 οδηγεί στα ακόλουθα βασικά συμπεράσματα:

1. Συμπληρωματικότητα και συγχρονισμός: Η FOMC και η ομοσπονδιακή κυβέρνηση ενέργησαν σε στενή συνεργασία, με τις νομισματικές παρεμβάσεις να παρέχουν άμεση ρευστότητα και τις δημοσιονομικές να εξασφαλίζουν άμεση οικονομική στήριξη.

2. Διαφορετικές χρονικές δυναμικές: Η νομισματική πολιτική μπορούσε να ανταποκριθεί ταχύτερα (μέσα σε ημέρες), ενώ η δημοσιονομική απαιτούσε περισσότερο χρόνο (εβδομάδες/μήνες) για νομοθετική διαδικασία.

3. Γλωσσική σταθερότητα έναντι Πρακτική προσαρμογή: Η FOMC διατήρησε έναν σταθερά προσεκτικό γλωσσικό τόνο ακόμη και καθώς οι πρακτικές της πολιτικής προσαρμόζονταν ριζικά, υπογραμμίζοντας τον ρόλο της ως εγγυητή μακροπρόθεσμης σταθερότητας.

4. Περιορισμένη προγνωστική αξία γλωσσικών δεικτών: Το συναίσθημα των πρακτικών αντανακλά περισσότερο τις τρέχουσες συνθήκες παρά προβλέπει μελλοντικές εξελίξεις, τουλάχιστον σε σχέση με την αγορά εργασίας.

Η πολυμεθοδική ανάλυση που παρουσιάστηκε στα κεφάλαια 4.1 έως 4.4 καταδεικνύει μια ιεραρχία κατανόησης: α) Στο επίπεδο λέξεων και προτάσεων (Sentiment Analysis), τα πρακτικά εκφράζουν έναν σταθερά προσεκτικό και κριτικό τόνο που κορυφώνεται στις περιόδους μεγαλύτερης κρίσης. β) Στο επίπεδο θεμάτων και προτεραιοτήτων (Topic Modeling), η FOMC μετατοπίζει το επίκεντρο της συζήτησής της ανάλογα με το

επικρατούσαν σοκ. γ) Στο επίπεδο πολιτικής δράσης και συντονισμού (ανάλυση πολιτικής ανταπόκρισης), αυτός ο γλωσσικός προβληματισμός και θεματική προσαρμογή ενσωματώνονται σε μια ευρύτερη, συγχρονισμένη στρατηγική διαχείρισης της μακροοικονομικής σταθερότητας. Τα γλωσσικά ευρήματα, επομένως, δεν είναι αποκομμένα σημεία δεδομένων, αλλά η επιφάνεια όψη μιας πολύπλοκης διαδικασίας λήψης αποφάσεων

ΚΕΦΑΛΑΙΟ 5: ΣΥΜΠΕΡΑΣΜΑΤΑ

Η παρούσα διπλωματική εργασία εξέτασε την επικοινωνία της Επιτροπής Ανοιχτής Αγοράς (FOMC) της Federal Reserve των ΗΠΑ κατά την περίοδο 2018-2024, η οποία περιλάμβανε τη σχετική ηρεμία των προ-πανδημικών ετών, το σεισμικό σοκ της πανδημίας COVID-19, την πληθωριστική έκρηξη, τις τραπεζικές αναταραχές και τις πρώτες ενδείξεις σταθεροποίησης. Χρησιμοποιώντας μια πολυμεθοδική προσέγγιση που συνδύασε τεχνικές Επεξεργασίας Φυσικής Γλώσσας (NLP), Ανάλυση Συναισθήματος (με λεξικό Loughran-McDonald και με το μοντέλο FinBERT) και Θεματική Ανάλυση (LDA) που πραγματοποίησε μια ποσοτική και ποιοτική διερεύνηση τόσο του περιεχομένου («τι») όσο και του συναισθηματικού τόνου («πως») στα Πρακτικά (Minutes) της FOMC. Παρακάτω συνοψίζονται τα κύρια ευρήματα της έρευνας.

5.1 ΚΥΡΙΑ ΕΥΡΗΜΑΤΑ ΤΗΣ ΕΡΕΥΝΑΣ

1. Ο συναισθηματικός τόνος ως ευαίσθητος δείκτης κρίσης:

Η ανάλυση επιβεβαιώνει ότι ο γλωσσικός τόνος των FOMC Minutes είναι σταθερά προσεκτικός και τείνει προς το αρνητικό, αποτελώντας έμφυτο χαρακτηριστικό της θεσμικής επικοινωνίας που στοχεύει στη διαχείριση κινδύνων. Οι διακυμάνσεις του τόνου είναι άμεσα συσχετισμένες με τα μακροοικονομικά σοκ. Η χρονική εξέλιξη των Sentiment Scores (τόσο του λεξικού όσο και του FinBERT) καταγράφει σαφείς κορυφώσεις αρνητικότητας:

a) Στην κορύφωση της πανδημίας COVID-19 (2020),

b) Στην έναρξη της πληθωριστικής κρίσης και του κύκλου σταθεροποίησης (2022) και
c) Κατά τις τραπεζικές αναταραχές(2023)

Η συγκριτική άσκηση μεταξύ των δύο μεθόδων ανάλυσης συναισθήματος ήταν καθοριστικής σημασίας. Το στατικό λεξικό Loughran-McDonald παρείχε διαφάνεια και επαναληψιμότητα, καταγράφοντας μια ουδέτερη προς αρνητική τάση. Το μοντέλο FinBERT, ικανό να κατανοεί το συμφραζόμενο, αποκάλυψε ένα βαθύτερο και πιο έντονα αρνητικό συναίσθημα, υποδεικνύοντας ότι η διαδικασία σκέψης της Επιτροπής επικεντρώνεται ουσιαστικά στην ανάλυση και την αντιμετώπιση κινδύνων. Αυτή η διαφορά δεν είναι αντίφαση, αλλά απόδειξη της συμπληρωματικότητας των μεθόδων και του βάθους που προσφέρει η κατανόηση συμφραζομένων.

2. Η χρονική εξέλιξη των θεματικών προτεραιοτήτων:

Η θεματική ανάλυση (LDA) σε δύο επίπεδα (παραγράφου και έτους) αποκάλυψε μια σταθερή και θεματική αρχιτεκτονική με πυλώνες όπως ο πληθωρισμός, η απασχόληση, οι χρηματοπιστωτικές συνθήκες και η γενικότερη οικονομική πολιτική. Η ετήσια ανάλυση LDA χαρτογράφησε μια δυναμική μετατόπιση προτεραιοτήτων που αντανακλούσε την άμεση ανταπόκριση της FOMC σε εξωγενείς σοκ, συγκεκριμένα:

a) Προ-πανδημική περίοδος (2018-2019): Εστίαση σε πληθωρισμό και ανάπτυξη
b) Περίοδος πανδημίας (2020-2021): Διάβρωση της κυριαρχίας του πληθωρισμού, άνοδος των θεμάτων σχετικά με χρηματοπιστωτικές αγορές και τη ρευστότητα και εισαγωγή της λέξης "pandemic" ως κεντρικό θέμα.
c) Πληθωριστική και Μετα-πανδημική περίοδος(2022-2023): Επιστροφή και ενίσχυση της συζήτησης για τον πληθωρισμό και τη νομισματική αυστηροποίηση, με την προσθήκη θέματος για τον τραπεζικό τομέα το 2023.
d) Περίοδος σταθεροποίησης (2024): Επιστροφή σε μια πιο ισορροπημένη θεματική κάλυψη, με αναβίωση του ενδιαφέροντος για την αγορά εργασίας ("labor market").

3. Η σύνδεση γλωσσικών ευρημάτων με την πολιτική πραγματικότητα:

Η έρευνα, πέραν της καταγραφής του συναισθήματος και των θεμάτων, επιχείρησε να συσχετίσει τα γλωσσικά ευρήματα με την πραγματική

πολιτική δράση και να αξιολογήσει την περιγραφική ή προγνωστική τους ισχύ. Τα αποτελέσματα αυτής της διερεύνησης οδηγούν σε δύο βασικές διαπιστώσεις:

a) Η γλώσσα ως άμεσος ανακλαστής της νομισματικής διαδικασίας: Η αντιστοίχιση των γλωσσικών δεικτών με ένα χρονολόγιο των πολιτικών παρεμβάσεων κατά την πανδημία COVID-19 (Κεφάλαιο 4.4) αποκάλυψε ότι τα ευρήματα δεν είναι αφηρημένα. Ο έκτακτα αρνητικός τόνος και η έντονη εμφάνιση θεμάτων που αφορούν τη ρευστότητα και τις χρηματοπιστωτικές αγορές το 2020 συνέπεσαν χρονικά και ήταν η άμεση λεκτική επιφάνεια των ίδιων των έκτακτων νομισματικών αποφάσεων. Οι λέξεις και οι προτάσεις στα πρακτικά λειτούργησαν ως ακριβής γλωσσικός καθρέφτης της δραματικής πολιτικής προσαρμογής (μηδενικά επιτόκια, απεριόριστη ποσοτική χαλάρωση).

b) Τα όρια της προγνωστικής ισχύος και η προτεραιότητα της κατανόησης: Η προσπάθεια να μετατραπεί αυτή η γλωσσική ανακλαστικότητα σε προγνωστικό εργαλείο απέδωσε περιορισμένα αποτελέσματα. Η στατιστική ανάλυση δεν αποκάλυψε ισχυρή συσχέτιση μεταξύ του συναισθήματος των πρακτικών και των μελλοντικών τιμών των εβδομαδιαίων αιτήσεων ανεργίας. Αυτό το κρίσιμο εύρημα υποδεικνύει ότι ο γλωσσικός τόνος λειτουργεί κυρίως ως δείκτης σύγχρονης αβεβαιότητας (coincident indicator). Δηλαδή, αποτυπώνει με ευαισθησία το συναισθηματικό και γνωστικό βάρος του παρόντος, αλλά δεν προσφέρει αξιόπιστη βάση για την πρόβλεψη συγκεκριμένων μελλοντικών οικονομικών αποτελεσμάτων. Η αξία του επομένως έγκειται στον διαφωτιστικό του ρόλο και στην αποκάλυψη του πλαισίου σκέψης, των προτεραιοτήτων και των αντιληπτών κινδύνων της Επιτροπής σε πραγματικό χρόνο, παρά σε μια ψευδο-μαθηματική προβλεπτική ικανότητα.

5.2 ΣΥΜΒΟΛΗ ΕΡΓΑΣΙΑΣ

Η παρούσα εργασία συνεισφέρει τόσο στην εμπειρική κατανόηση της νομισματικής επικοινωνίας όσο και στη μεθοδολογική συζήτηση. Παρέχει μια ολιστική και διαχρονική αφήγηση για το πώς η FOMC επικοινώνησε και προσάρμοσε τις προτεραιότητές της μέσα από διαδοχικές κρίσεις. Επαληθεύει ότι τα Minutes είναι πολύ περισσότερα από απλά τεχνικά

αρχεία και αποτελούν ζωντανά έγγραφα που καταγράφουν τη διαλεκτική σχέση μεταξύ γλώσσας, πολιτικής και οικονομικής πραγματικότητας. Επιπλέον, υλοποιεί και συγκρίνει συστηματικά δύο διαφορετικές προσεγγίσεις Ανάλυσης Συναισθήματος (λεξικολογική και νευρωνικά δίκτυα) στον ίδιο πυρήνα δεδομένων, επιδεικνύοντας τις διακριτικές τους δυνατότητες και τον τρόπο συμπλήρωσής τους. Η εφαρμογή της LDA σε δύο επίπεδα ανάλυσης (παράγραφος και έτος) προσφέρει έναν ιεραρχικό φακό για την κατανόηση τόσο των διαχρονικών μοτίβων όσο και των χρονικών μετατοπίσεων.

Παρά τις προσεκτικές μεθοδολογικές επιλογές, η έρευνα έχει κάποιους περιορισμούς που ανοίγουν το δρόμο για μελλοντική εξερεύνηση: Αρχικά, συναντιούνται περιορισμοί στη Γλωσσική Ανάλυση, καθώς και το FinBERT, αν και προηγμένο, έχει το δικό του όριο στη κατανόηση της πολύ εξειδικευμένης γλώσσας των κεντρικών τραπεζών.Επίσης, υπάρχει περιορισμός στα δεδομένα, καθώς η ανάλυση περιορίστηκε στα επίσημα Πρακτικά (Minutes). Η συμπερίληψη και άλλων μορφών επικοινωνίας (ομιλίες μελών, συνεντεύξεις τύπου, δημοσιεύματα) θα μπορούσε να προσφέρει μια πιο πλούσια εικόνα της επικοινωνιακής στρατηγικής.

Συμπερασματικά αυτό που αξίζει κάποιος να κρατήσει από αυτή την εργασία είναι ότι η πορεία από τον προ-πανδημικό προβληματισμό για τον πληθωρισμό, στην έκτακτη διαχείριση της πανδημικής κρίσης, στη μάχη ενάντια στην πληθωριστική έκρηξη και τέλος στην αναζήτηση σταθεροποίησης, δεν είναι μόνο μια ιστορία πολιτικών αποφάσεων, αλλά και μια ιστορία γλώσσας. Τα FOMC Minutes αποτελούν τον σημαντικότερο δημόσιο αρχείο αυτής της γλωσσικής πορείας. Η παρούσα έρευνα έδειξε ότι μέσω της συστηματικής ανάλυσης αυτών των κειμένων, μπορούμε να αποκαλύψουμε όχι μόνο τι απασχολεί τους πολιτικούς, αλλά και πώς σκέφτονται και αντιλαμβάνονται τον κόσμο γύρω τους σε στιγμές έντονης αβεβαιότητας. Σε μια εποχή όπου η επικοινωνία των κεντρικών τραπεζών έχει γίνει ένα ισχυρό εργαλείο πολιτικής, η κατανόηση της γλώσσας τους παραμένει βασική για την κατανόηση της οικονομικής μας πραγματικότητας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Araci, D. (2019). FinBERT: Financial sentiment analysis with pre-trained language models (arXiv:1908.10063). arXiv. http://arxiv.org/abs/1908.10063

Bernanke, B. S. (2020). *The new tools of monetary policy. American Economic Review, 110*(4), 943–983. https://doi.org/10.1257/aer.110.4.943

Bird, S., Klein, E., & Loper, E. (2009). *Natural language processing with Python.* O'Reilly Media, Inc.

Blei, D. M. (2012). *Probabilistic topic models. Communications of the ACM, 55(*4), 77–84. https://doi.org/10.1145/2133806.2133826

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, *3*(Jan), 993–1022. https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf

Board of Governors of the Federal Reserve System. (2020). Federal Reserve announces extensive new measures to support the economy. https://www.federalreserve.gov/newsevents/pressreleases/monetary20200323a.htm

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics*: Human Language Technologies, Volume 1 (Long and Short Papers), 4171–4186. https://doi.org/10.18653/v1/N19-1423

Federal Reserve. (n.d.). *Federal Open Market Committee (FOMC) minutes*. Retrieved November 26, 2025, from https://www.federalreserve.gov/monetarypolicy/fomccalendars.htm

Field, A. (2018). Discovering statistics using IBM SPSS statistics (5th ed.). Sage.

Gentzkow, M., Kelly, B., & Taddy, M. (2019). Text as data. *Journal of Economic Literature*, *57*(3), 535–574. https://doi.org/10.1257/jel.20181020

Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics, 133*(2), 801–870. https://doi.org/10.1093/qje/qjx045

Jurafsky, D., & Martin, J. H. (2023). *Speech and language processing* (3rd ed.). Pearson

Loughran, T., & McDonald, B. (2011). When is a liability not a liability? Textual analysis, dictionaries, and 10-Ks. *The Journal of Finance, 66*(1), 35–65. https://doi.org/10.1111/j.1540-6261.2010.01625.x

Mishkin, F. S. (2022). The economics of money, banking, and financial markets (13th ed.). Pearson.

Rehurek, R., & Sojka, P. (2010). Software framework for topic modelling with large corpora. *In Proceedings of the LREC 2010 workshop on new challenges for NLP frameworks* (pp. 45–50). ELRA.

Shin, D. (2021). Topic modeling of FOMC minutes: Evolution of discussions before and after the COVID-19 outbreak. *Finance Research Letters, 43*, 101995. https://doi.org/10.1016/j.frl.2021.101995

U.S. Bureau of Labor Statistics (BLS). (n.d.). Labor Force Statistics (CPS), Unemployment Rate [Data set]

U.S. Department of the Treasury. (2020). The CARES Act Works for All Americans.

Wolf, T., Debut, L., Sanh, V., Chaumond, J., Delangue, C., Moi, A., Cistac, P., Rault, T., Louf, R., Funtowicz, M., Davison, J., Shleifer, S., von Platen, P., Ma, C., Jernite, Y., Plu, J., Xu, C., Le Scao, T., … Rush, A. M. (2020). Transformers: State-of-the-art natural language processing. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* (pp. 38–45). Association for Computational Linguistics. https://doi.org/10.18653/v1/2020.emnlp-demos