ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΑΝΘΡΩΠΙΣΤΙΚΩΝ ΚΑΙ ΚΟΙΝΩΝΙΚΩΝ ΕΠΙΣΤΗΜΩΝ ΚΑΙ ΔΙΚΑΙΟΥ
ΕΡΓΑΣΤΗΡΙΟ ΘΕΩΡΗΤΙΚΗΣ ΚΑΙ ΕΦΑΡΜΟΣΜΕΝΗΣ ΟΙΚΟΝΟΜΙΚΗΣ ΚΑΙ ΔΙΚΑΙΟΥ

Στατιστική ανάλυση των παραγόντων που επηρεάζουν το πλήθος νικών μιας ομάδας μπάσκετ του NBA

Διπλωματική Εργασία

της

Αργυρούς Δαμουλάκη

Επίβλεψη: Παναγιώτης Μιχαηλίδης, Αν. Καθηγητής ΕΜΠ

Αθήνα, 2021

Πρόλογος

Η παρούσα διπλωματική εργασία εκπονήθηκε στο πλαίσιο των προπτυχιακών μου σπουδών στη Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου και συγκεκριμένα στον τομέα Ανθρωπιστικών και Κοινωνικών Επιστημών και Δικαίου. Αρχικά, θα ήθελα να ευχαριστήσω τον Αναπληρωτή Καθηγητή ΕΜΠ, κύριο Μιχαηλίδη Παναγιώτη, για την ανάθεση του θέματος της διπλωματικής εργασίας, με την οποία μου έδωσε την ευκαιρία να ασχοληθώ με ένα αντικείμενο που επιθυμούσα πάνω στην αθλητική στατιστική. Εν συνεχεία, θα ήθελα να ευχαριστήσω τον Καθηγητή ΕΜΠ κύριο Τσώλα Ιωάννη και το μέλος Ε.ΔΙ.Π. ΕΜΠ κύριο Στέλιο Σπυρίδων για τη συμμετοχή τους στην τριμελή επιτροπή. Τέλος, ιδιαίτερες ευχαριστίες οφείλω στον Διδάκτορα ΕΜΠ κύριο Μαρίνο Θεοχάρη για την καθοδήγηση, την εμπιστοσύνη και τη διάθεσή του καθ' όλη τη διάρκεια εκπόνησης της παρούσας διπλωματικής, καθώς επίσης και στο συνάδελφο Μυλωνά Διονύση.

Περίληψη

Στόχος της παρούσας διπλωματικής εργασίας είναι η μελέτη των παραγόντων που επιδρούν κατά κύριο λόγο στο πλήθος νικών της κανονικής περιόδου που πετυχαίνει μια ομάδα καλαθοσφαίρισης στο NBA. Για τον σκοπό της μελέτης, θα εκτιμηθεί ένα Δυναμικό Χωρικό Υπόδειγμα Durbin (Dynamic Spatial Durbin Model).

Πρωτοπόροι στην ανάλυση στατιστικών και παραγόντων που οδηγούν μια ομάδα σε νίκη είναι ο Dean Oliver (2004) και John Hollinger (2002, 2003, 2004 και 2005). Ο πρώτος καθόρισε τους τέσσερις παράγοντες (4 factors) οι οποίοι παρακίνησαν και συνεχίζουν να παρακινούν αρκετούς ερευνητές να ασχοληθούν με τον συγκεκριμένο τομέα. Ο δεύτερος ανέπτυξε κάποιους αγωνιστικούς δείκτες και, ειδικότερα, είναι ευρέως γνωστός για τον δείκτη που ποσοτικοποιεί την επίδοση ενός παίκτη (PER) και έχει καθιερωθεί στην αθλητική στατιστική στην καλαθοσφαίριση.

Οι παράγοντες που μελετώνται παρακάτω, όμως, δεν είναι εξ' ολοκλήρου αγωνιστικοί. Στο συγκεκριμένο πρωτάθλημα έχουν τεθεί κανονισμοί που το καθιστούν όσο γίνεται πιο «δίκαιο» για τις αθλητικές επιχειρήσεις. «Δίκαιο» με την έννοια ότι όλες οι ομάδες έχουν ένα μέγιστο οικονομικό όριο δαπανών (salary cap) κι έναν φόρο πολυτελείας στον οποίο υπάγονται αν ξεπεράσουν κατά πολύ το salary cap. Επίσης, υπάρχουν κανόνες που καθιστούν τις ανταλλαγές μεταξύ των ομάδων ισοδύναμες. Κατά τη διαδικασία επιλογής νεοσύλλεκτων παικτών (draft), κάθε σεζόν πριν την έναρξη της αγωνιστικής περιόδου, οι ομάδες που τερμάτισαν χαμηλότερα στη βαθμολογία έχουν το πλεονέκτημα να διαλέξουν στις πρώτες θέσεις, βάζοντας στο στόχαστρο τους καλύτερους διαθέσιμους παίκτες ή και τους περισσότερο ταιριαστούς στο ρόστερ τους. Αναπόφευκτα, λοιπόν, η αγωνιστική επίδοση των ομάδων είναι αποτέλεσμα και των διοικητικών αποφάσεων σχετικά με τις ανταλλαγές, τη διαχείριση των νεοσύλλεκτων παικτών που επιλέγει κάθε ομάδα και, φυσικά, το ύψος των μισθών που είναι διατεθειμένη να προσφέρει στο δυναμικό της.

Μια ομάδα, βέβαια, δεν απασχολεί μόνο αθλητές, προπονητές και αναλυτές αλλά ακόμα τον γενικό διευθυντή της ομάδας και παράγοντες της διοίκησης, ανθρώπους

στον τομέα της επικοινωνίας, των οικονομικών, των πωλήσεων, της διαφήμισης, ιατρικό προσωπικό, νομικούς κλπ. Γενικά, η λίγκα από μόνη της αποτελεί μια οικονομική μονάδα και όλα καθορίζονται από μία συλλογική σύμβαση. Στο πλαίσιο αυτό, οι ομάδες συνιστούν «επιχειρήσεις» με ιδιαίτερο χαρακτήρα. Ως εκ τούτου, μεγάλες επιχειρήσεις συνεργάζονται με τη λίγκα, κάθε ομάδα ξεχωριστά ή ακόμα και με ορισμένους παίκτες ατομικά. Πλέον, στη σύγχρονη εποχή του NBA, τα χρηματικά ποσά που καλούνται οι ομάδες να διαχειριστούν είναι τεράστια. Ως αποτέλεσμα, η ίδια η διοργάνωση αποτελεί ένα ελκυστικό αντικείμενο μελέτης και ακόμα περισσότερο η επίδραση των αποφάσεων κάθε ομάδας με την αγωνιστική της επίδοση.

Συνεπώς, στην παρούσα εργασία εκτιμάται κατά πόσο οι σχετικές αποφάσεις μιας ομάδας επηρεάζουν την αγωνιστική επίδοση των υπολοίπων. Θα δούμε, μάλιστα, ότι το ύψος των μισθών είναι αυτό που επηρεάζει περισσότερο αγωνιστικά τις υπόλοιπες ομάδες. Τέλος, θα δούμε κατά πόσο η αγωνιστική επίδοση της προηγούμενης περιόδου, τόσο της ίδιας της ομάδας όσο και των αντιπάλων της, επηρεάζει αγωνιστικά την επόμενη σεζόν.

Abstract

The aim of this thesis is to study the factors that, primarily, affect the number of regular season wins achieved by an NBA basketball team. For the purpose of the study, a Dynamic Spatial Durbin Model will be estimated.

Pioneers in analyzing statistics and factors that lead a team to win are Dean Oliver (2004) and John Hollinger (2002, 2003, 2004 and 2005). The first, established the “Four Factors to win” that motivated and still motivate several researchers. The second, developed some basketball statistics and is widely known for the rate that quantifies the performance of a player (PER) which has been established in basketball statistics.

However, the factors studied below are not entirely related to game statistics. There are regulations to make the league as “fair” as possible for the teams. “Fair” in the sense that all groups have a maximum financial spending limit (salary cap) and a luxury tax to which they fall, if they far exceed salary cap. Also, there are rules in order to secure that trades between NBA teams are equivalent. In the process of draft selection, before the beginning of the season, teams that finished lower in the ratings have the advantage of picking playerσ in the first places, targeting to the best available players or the most suited to their roster. Inevitably, therefore, the game performance of the teams is a result of decisions regarding the trades, the management of the rookie players and, of course, the amount of wages it is willing to offer to its players.

A team employs not only athletes, coaches, analysts and statisticians but still general managers and management agents, people in the fields of communication, finance, sales, advertising, medical staff, legal, etc. In general, the league is an economic unit, and everything is determined by agreements. Within this context, the teams are "enterprises" with a particular commercial character. Therefore, large companies work with the league, each team individually or even with certain players individually. Now, in the modern era of the NBA, the amounts of money that teams are called to manage are huge. As a result, the NBA league itself is an attractive

object of study and even more the effect of each team decisions on the game performance.

Therefore, in this work we estimate whether the decisions of a team affect the game performance of other teams. We will see that the payroll of a team is what affects its contenders most. Finally, we will examine whether the previous season’s game performance of a team affects next season's performance, not only for the team itself but also for its opponents.

Πίνακας περιεχομένων

Κεφάλαιο 1: Εισαγωγή **1**

1.1 Εισαγωγικά στοιχεία 1

1.2 Τρόπος διεξαγωγής πρωταθλήματος 5

1.2.1 Διαχωρισμός των ομάδων 5

1.2.2 Κανονική Περίοδος (Regural Season) 6

1.2.3 Playoffs 7

1.3 Ύψος Συμβολαίων, Φόρος Πολυτελείας και Πολιτική Συμβολαίων 8

1.4 NBA Draft 10

1.5 Ανταλλαγές (Trades) 11

1.6 Αντικείμενο, στόχοι και δομή εργασίας 11

Κεφάλαιο 2: Βιβλιογραφική Ανασκόπηση **13**

2.1 Αθλητική Στατιστική 13

2.2 Χωρικά Οικονομετρικά Υποδείγματα 23

2.3 Συνεισφορά στη βιβλιογραφία 30

Κεφάλαιο 3: Μεθοδολογικό πλαίσιο **32**

3.1 Γενικευμένο υπόδειγμα 32

3.1.1 Το Υπόδειγμα Χωρικής Διάχυσης Durbin (Spatial Durbin Model) 32

3.1.2 Το Δυναμικό Υπόδειγμα Χωρικής Διάχυσης Durbin (Dynamic Spatial Durbin Model ή D-SDM) 34

3.1.3 Άμεσες, Έμμεσες και Συνολικές Επιδράσεις 35

3.1.4 Η Μέθοδος της Μέγιστης Πιθανοφάνειας σε Δεδομένα Panel 36

3.2 Σχετικοί έλεγχοι 41

3.2.1 Επιλογή υποδείγματος σταθερών ή τυχαίων επιδράσεων 41

3.2.2 Έλεγχος στασιμότητας (Stationarity test) 42

3.2.3 Έλεγχος μη-αιτιότητας Granger (No-Causality test) 43

3.2.4 Έλεγχος Ευστάθειας Συστήματος (Stability test).....45
3.3 Μελέτη Περίπτωσης.....46
3.3.1 Εξειδίκευση του υποδείγματος 46
3.3.2 Χωρική Μήτρα βαρών 48

Κεφάλαιο 4: Εφαρμογή.....**50**

4.1 Δεδομένα και πηγές δεδομένων.....50
4.2 Εμπειρικά αποτελέσματα 53

Κεφάλαιο 5: Συμπεράσματα **59**

5.1 Συμπεράσματα 59

Βιβλιογραφία **64**

Παραρτήματα.....**70**
ΠΑΡΑΡΤΗΜΑ Ι: Δεδομένα Πάνελ (Panel Data) 70
Ι1. Έννοια.....70
Ι2. Το βασικό υπόδειγμα με δεδομένα panel 70
Ι3. Μοντέλο Σταθερών Επιδράσεων (FEM - Fixed Effect Model) 72
Ι4. Μοντέλο Τυχαίων Επιδράσεων (REM - Random Effect Model).....72
Ι5. Υπόδειγμα FEM ή REM.....73
Ι6. Δυναμικά Υποδείγματα με Δεδομένα Panel (Dynamic Panel Data Model–DPDM) 75
Ι7. Έλεγχος Μοναδιαίας Ρίζας 75
ΠΑΡΑΡΤΗΜΑ ΙΙ: Επαυξημένος Έλεγχος Μοναδιαίας Ριζας των Dickey-Fuller (ADF Test) 77
ΠΑΡΑΡΤΗΜΑ ΙΙΙ: Προηγμένα Στατιστικά (Advanced Stats) 79
ΠΑΡΑΡΤΗΜΑ IV: Οι ομάδες του NBA 81
ΠΑΡΑΡΤΗΜΑ V: Μήτρα Βαρών (W) και Κανονικοποιημένη Μήτρα Βαρών (Wnorm) 82

Ευρετήριο Πινάκων

Κεφάλαιο 1 ..... **1**

Πίνακας 1.1: Διαχωρισμός των ομάδων του NBA 6

Πίνακας 1.2: Επίπεδο φόρων (Tax levels) από την αγωνιστική περίοδο 2013-2014 9

Κεφάλαιο 4 ..... **50**

Πίνακας 4.1: Αποτελέσματα LLC (Levin, Lin και Chu) Unit Root Test 53

Πίνακας 4.2: Αποτελέσματα Dumitrescu and Hurlin (2012) Non-Causality Test 54

Πίνακας 4.3: Περιγραφικά Στατιστικά 55

Πίνακας 4.4: Περιγραφικά Στατιστικά με φυσικό λογάριθμο 55

Πίνακας 4.5: Αποτελέσματα του Δυναμικού Υποδείγματος Χωρικής Διάχυσης Durbin 56

Παραρτήματα **70**

Πίνακας Π1: Ομάδες NBA 81

Πίνακας Π2: Μήτρα Βαρών 82

Πίνακας Π3: Κανονικοποιημένη Μήτρα Βαρών 83

Κεφάλαιο 1: Εισαγωγή

1.1 Εισαγωγικά στοιχεία

Το NBA (National Basketball Association) είναι η διοργανώτρια αρχή του ομώνυμου πρωταθλήματος καλαθοσφαίρισης των ΗΠΑ. Απαρτίζεται από 30 ομάδες (29 με έδρα στις ΗΠΑ και 1 με έδρα στον Καναδά) και είναι μέλος της Ομοσπονδίας Καλαθοσφαίρισης των ΗΠΑ (USAB), η οποία είναι αναγνωρισμένη από τη Διεθνή Ομοσπονδία Καλαθοσφαίρισης (FIBA) ως αρμόδια για το άθλημα στις Ηνωμένες Πολιτείες. Η λίγκα ιδρύθηκε στη Νέα Υόρκη στις 6 Ιουνίου του 1946 με την επωνυμία BAA (Basketball Association of America). Στις 3 Αυγούστου του 1949 μετονομάστηκε σε NBA (National Basketball Association), έπειτα από τη συγχώνευση της BAA με την ανταγωνίστριά της NBL (National Basketball League)[1].

Αρχικά, το NBA ήταν ένα πρωτάθλημα δυσφημισμένο στις Ηνωμένες Πολιτείες. Το 1984, η αγοραία αξία του NBA μειώθηκε και 17 από τις 23 ομάδες ήταν στα πρόθυρα της χρεοκοπίας. Την ίδια χρονιά ο David Stern ανέλαβε τη θέση του «Κομισάριου» και εισήγαγε μια νέα εποχή για το πρωτάθλημα, μετατρέποντάς το σε εμπορικό μεγαθήριο που μεταδίδεται τώρα σε 212 χώρες σε 42 γλώσσες.

Πλέον, το πρωτάθλημα του NBA συνιστά ένα από τα τέσσερα μεγάλα πρωταθλήματα σε επίπεδο επαγγελματικού αθλητισμού στη Βόρεια Αμερική (baseball, football, American football). Ως αθλητικό γεγονός, το μπάσκετ έχει πολύ μεγάλη ζήτηση σε ένα περιβάλλον υψηλού ανταγωνισμού, αν αναλογιστούμε τα υπόλοιπα ομαδικά αθλήματα και τον τεράστιο πληθυσμό της Αμερικής, όπου διεξάγεται το πρωτάθλημα. Σύμφωνα με το SPORTS SHOW[2] και το STATSCORE[3], πρόκειται για την τρίτη πιο διάσημη αθλητική διοργάνωση στην Αμερική το 2021 αλλά και παγκοσμίως με βάση το TOTAL SPORTEK[4].

---

[1] Πηγή: The Date in the NBA: https://www.nba.com/history/this-date-aug
[2] https://sportsshow.net/most-popular-sports-in-america/
[3] https://blog.statscore.com/what-are-the-5-most-popular-sports-in-the-usa/
[4] https://www.totalsportek.com/worlds-popular-sports-eleven/

Ο χαρακτήρας και η οργάνωση του πρωταθλήματος βασίζονται στο επιχείρημα ότι αν υπήρχαν πολλά επιμέρους εθνικά πρωταθλήματα και όχι ένα, θα μοιράζονταν σε αυτά οι καλύτεροι παίκτες και, συνεπώς, το κάθε ένα από αυτά θα ήταν πιθανώς λιγότερο ανταγωνιστικό και, συνεπώς, ευχάριστο για να το παρακολουθήσει κάποιος. Έτσι, όχι μόνο θα προσελκύονταν λιγότεροι θεατές αλλά οι ομάδες και κάθε επιμέρους ένωση θα προσκόμιζαν λιγότερα κέρδη. Αντίθετα, ένα ισχυρό πρωτάθλημα μονοπωλεί όλους τους ποιοτικούς παίκτες και επεκτείνεται για να διατηρήσει ένα υψηλό επίπεδο ανταγωνισμού, με αποτέλεσμα κάθε παιχνίδι να είναι πιο ευχάριστο και ελκυστικό για το «αγοραστικό» κοινό. Ο James J. Zhang (2005), *Importance of In-Game Entertainment Amenities at Professional Sporting Events: A Case for NBA Season Ticket Holders*[5], αναφέρει ότι στην πρώιμη ιστορία του επαγγελματικού μπάσκετ, τα παιχνίδια δεν ήταν τίποτα περισσότερο από τους θεατές που πήγαιναν σε αρένες και παρακολουθούσαν τους παίκτες δύο ομάδων να ανταγωνίζονται. Καθώς ο χρόνος έχει περάσει και η αθλητική βιομηχανία στο σύνολό της έχει εξελιχθεί, τα επαγγελματικά παιχνίδια μπάσκετ έχουν εξελιχθεί σε ένα γεγονός ψυχαγωγίας. Ωστόσο, σε αντίθεση με άλλες μορφές δημοφιλών τμημάτων της βιομηχανίας ψυχαγωγίας (όπως ταινίες και μουσική), τα παιχνίδια μπάσκετ έχουν διαλείμματα μεταξύ των ενεργειών. Αν και οι αθλητές, οι προπονητές και τα ταλέντα τους είναι το βασικό προϊόν, σύμφωνα με τους Mullin, Hardy και Sutton (2000), είναι ένα δύσκολο ζήτημα για τους managers του αθλητισμού να διασκεδάζουν τους θεατές σε μια εκδήλωση για να μεγιστοποιήσουν τα ενδιαφέροντά τους, ιδιαίτερα όταν το ίδιο το παιχνίδι δεν είναι ελκυστικό. Πάνω σε αυτή τη γραμμή, το NBA έχει εξελιχθεί σε ένα τεράστιο επιχειρηματικό πρωτάθλημα στη διαδικασία ανάπτυξης, επέκτασης και μονοπώλησης.

Εσωτερικά, ως ανώτατη αρχή του NBA, το Διοικητικό Συμβούλιο λαμβάνει αποφάσεις για όλα τα θέματα κι έχει εργασιακή σχέση με τον Κομισάριο, ο οποίος αποτελεί εκτελεστικό όργανο. Οι ομάδες με τη σειρά τους απλώς συμμορφώνονται με τις αποφάσεις που λαμβάνονται από το Διοικητικό Συμβούλιο.

[5] https://www.researchgate.net/publication/292706398_Importance_of_ingame_entertainment_amenities_at_professional_sporting_events_A_case_for_NBA_season_ticket_holders.

Εξωτερικά, όλες οι ομάδες στο NBA διαπραγματεύονται συλλογικά για ό,τι αφορά το επιχειρηματικό κομμάτι. Επειδή το περιεχόμενο του παιχνιδιού προστατεύεται από τον νόμο περί πνευματικών δικαιωμάτων, οι ραδιοτηλεοπτικοί φορείς πρέπει να πληρώσουν το πρωτάθλημα για τα δικαιώματα μετάδοσης των αγώνων. Με τη διαπραγμάτευση με τους ραδιοφωνικούς σταθμούς ως ενοποιημένος σύλλογος έχουν αρκετά υψηλά κέρδη μετάδοσης.

Οι παίκτες που αγωνίζονται στο πρωτάθλημα του NBA ανήκουν στους καλύτερα αμειβόμενους αθλητές του κόσμου. Ενδεικτικά, στη λίστα των πενήντα καλύτερα αμειβόμενων αθλητών παγκοσμίως του Forbes[6] υπάρχουν δεκατρείς παίκτες του NBA από τους οποίους οι επτά βρίσκονται στην πρώτη εικοσάδα. Στην καθημερινότητα του NBA γίνεται λόγος για συμβόλαια τεράστιων ποσών και είναι λογικό αν αναλογιστούμε τα υπέρογκα έσοδα που λαμβάνει η συγκεκριμένη διοργάνωση. Για παράδειγμα την τελευταία χρονιά που μελετάμε 2019-2020 οι ομάδες του NBA είχαν συνολικά έσοδα ύψους περίπου 7,92 δις δολάρια σύμφωνα με το STATISTA[7], ενώ την προηγούμενη χρονιά, 2018-2019, τα συνολικά έσοδα ήταν ακόμα υψηλότερα στο ύψος των 8,76 δις δολαρίων.

Οι πηγές εισοδήματος της Ένωσης του NBA που σχετίζονται άμεσα με τις ομάδες προέρχονται από την αγορά εισιτηρίων και εντός γηπέδου προϊόντων ή υπηρεσιών, τα τηλεοπτικά συμβόλαια με τα κανάλια μετάδοσης των αγώνων, τα διαφημιστικά λογότυπα-σήματα (logo) στις φανέλες των παικτών αλλά και τις πωλήσεις προϊόντων με το σήμα του NBA. Πρόκειται, λοιπόν, για μπασκετικά εμπορικά έσοδα, Basketball Related Income (BRI). Έσοδα που δεν περιλαμβάνονται στα BRI είναι αυτά από τα συμβόλαια και τα πρόστιμα παικτών, τις ασφάλειες, την αγορά ή διάθεση ακινήτων και περιουσιακών στοιχείων, τις φιλανθρωπίες και φυσικά τα έσοδα των ομάδων (revenue sharing).

Είναι σημαντικό ότι παρά τα τεράστια ποσά εσόδων, οι άνθρωποι της λίγκας δεν επαναπαύονται. Παρατηρείται ότι όσο περισσότερα είναι τα έσοδα τόσο πιο ποιοτικό είναι το πρωτάθλημα προσφέροντας καλύτερο θέαμα. Ας μην ξεχνάμε ότι πρόκειται για ένα μέσο ψυχαγωγίας το οποίο αναβαθμίζεται συνεχώς και

[6] https://www.forbes.com/athletes/
[7] https://www.statista.com/statistics/193467/total-league-revenue-of-the-nba-since-2005/

προσελκύει όλο και περισσότερο κόσμο. Σύμφωνα με το National Post[8], το 2021 παρατηρήθηκε αύξηση ύψους 15% των NBA league pass στην Ευρώπη για την παρακολούθηση του πρωταθλήματος, το οποίο πλέον είναι διαθέσιμο σε περισσότερες από 200 χώρες παγκοσμίως. Ως εκ τούτου, πραγματοποιούνται πολλές διαφημιστικές καμπάνιες σε όλο τον κόσμο προβάλλοντας το πρωτάθλημα και ιδιαίτερα όταν πλησιάζει η περίοδος του All Star Game ή των playoffs. Επίσης, πωλούνται προϊόντα με το logo του NBA, ομάδων ή ακόμα και παικτών, σε συνεργασία με ισχυρές επιχειρήσεις. Εξάλλου, από το 1990-1991 διοργανώνονται διεθνείς αγώνες ξεκινώντας από την Ιαπωνία, το Μεξικό και, έπειτα, τη Δυτική Ευρώπη. Τέλος, αξίζει να αναφέρουμε ότι πλέον το NBA έχει επεκταθεί στην ευρωπαϊκή αγορά και ειδικότερα με τη σημαντική απορρόφηση Ευρωπαίων παικτών στο πρωτάθλημα. Για παράδειγμα, σύμφωνα με το National Post[9], το 2021 στην Ιταλία, Ισπανία και Μεγάλη Βρετανία η τηλεθέαση του NBA αυξήθηκε κατά ένα σημαντικό ποσοστό της τάξεως του 32%, 17% και 72%, αντιστοίχως, με την άνοδο ευρωπαίων παικτών όπως οι Antetokounmpo, Doncic και Jokic. Μάλιστα, η ραγδαία βελτίωση του τελευταίου συνέβαλε στην αύξηση της τηλεθέασης του NBA ύψους 200% στη χώρα του, τη Σερβία.

Συνοψίζοντας, ο εμπορικός και επιχειρηματικός χαρακτήρας του NBA καθιστά τη διοργάνωση ένα ιδιαίτερα ενδιαφέρον αντικείμενο μελέτης. Όπως και σε μία επιχείρηση στόχος είναι η μεγιστοποίηση των κερδών, έτσι και για κάθε επιμέρους ομάδα του πρωταθλήματος είναι η μεγιστοποίηση των νικών της, που συνεπάγεται αύξηση του ανταγωνισμού και, κατ' επέκταση, περισσότερο κοινό.

Εκ πρώτης όψεως, η νίκη ενός παιχνιδιού φαίνεται να επηρεάζεται μόνο από αγωνιστικούς παράγοντες. Ωστόσο, πίσω από την αγωνιστική εικόνα κάθε ομάδας «κρύβεται» μια σειρά αποφάσεων και ενεργειών της διοίκησης που σχετίζεται με τον οικονομικό (συγκεκριμένα το ύψος των συμβολαίων των παικτών) καθώς και τον αγωνιστικό τομέα (όπως επιλογή του προπονητικού επιτελείου, των παικτών κλπ.).

[8] https://nationalpost.com/pmn/sports-pmn/european-talent-powers-overseas-ratings-boom.
[9] https://nationalpost.com/pmn/sports-pmn/european-talent-powers-overseas-ratings-boom.

Προκειμένου να μελετηθούν οι παράγοντες που επηρεάζουν το ποσοστό νικών μιας ομάδας στην παρούσα εργασία, κρίνεται χρήσιμο να παρουσιαστούν στη συνέχεια ο τρόπος διεξαγωγής του πρωταθλήματος και ορισμένοι βασικοί κανονισμοί που εφαρμόζονται με στόχο μία λίγκα ανταγωνιστική και δίκαιη.

1.2 Τρόπος διεξαγωγής πρωταθλήματος[10]

1.2.1 Διαχωρισμός των ομάδων

Για την καλύτερη διεξαγωγή του πρωταθλήματος οι ομάδες έχουν χωριστεί σε δύο περιφέρειες (conference), την Ανατολική και τη Δυτική. Αυτές με την σειρά τους χωρίζονται σε τρία μικρότερα γκρουπ ομάδων (division). Η Ανατολική Περιφέρεια χωρίζεται στις κατηγορίες Atlantic (Ατλαντική), Central (Κεντρική) και Southeast (Νοτιοανατολική), ενώ η Δυτική στις Southwest (Νοτιοδυτική), Northwest (Βορειοδυτική) και Pacific (Ειρηνικού) (Πίνακας 1.1).

[10] Πηγές: Wikipedia (https://en.wikipedia.org/wiki/National_Basketball_Association)

Πίνακας 1.1: Διαχωρισμός των ομάδων του NBA

Περιφέρεια (Conference)			
ΑΝΑΤΟΛΗ		**ΔΥΣΗ**	
Κατηγορία (Division)	**Ομάδα**	**Κατηγορία (Division)**	**Ομάδα**
Atlantic	Boston Celtics	Southwest	Dallas Mavericks
	Brooklyn Nets		Houston Rockets
	New York Knicks		Memphis Grizzlies
	Philadelphia 76ers		New Orleans Pelicans
	Toronto Raptors		San Antonio Spurs
Central	Chicago Bulls	Northwest	Denver Nuggets
	Cleveland Cavaliers		Minnesota Timberwolves
	Detroit Pistons		Portland Trail Blazers
	Indiana Pacers		Oklahoma City Thunder
	Milwaukee Bucks		Utah Jazz
Southeast	Atlanta Hawks	Pacific	Golden State Warriors
	Charlotte Hornets		Los Angeles Clippers
	Miami Heat		Los Angeles Lakers
	Orlando Magic		Phoenix Suns
	Washington Wizards		Sacramento Kings

1.2.2 Κανονική Περίοδος (Regural Season)

Η κανονική σεζόν του NBA ξεκινά την τελευταία εβδομάδα του Οκτωβρίου. Κατά τη διάρκεια της κανονικής περιόδου, κάθε ομάδα παίζει 82 παιχνίδια, 41 εντός και 41 εκτός έδρας. Μια ομάδα αντιμετωπίζει αντιπάλους στη δική της κατηγορία (division) τέσσερις φορές (16 παιχνίδια, οχτώ εντός έδρας και οχτώ εκτός). Κάθε ομάδα παίζει με έξι από τις ομάδες των υπόλοιπων δύο divisions της περιφέρειάς της τέσσερις φορές (24 παιχνίδια) και με τις υπόλοιπες τέσσερις ομάδες τρεις φορές

(12 παιχνίδια). Τέλος, κάθε ομάδα παίζει με όλες τις ομάδες της άλλης περιφέρειας δύο φορές, και συνεπώς, πρόκειται στο σύνολο για 30 παιχνίδια με την άλλη περιφέρεια.

Τον Φεβρουάριο η κανονική περίοδος παύει, για να πραγματοποιηθεί το ετήσιο NBA All-Star Game. Λίγο μετά τη διακοπή του πρωταθλήματος για το All-Star ακολουθεί η περίοδος των ανταλλαγών με συγκεκριμένη προθεσμία. Οι μεγάλες ανταλλαγές συνήθως ολοκληρώνονται ακριβώς πριν από την προθεσμία για τις ανταλλαγές (trade deadline).

Γύρω στα μέσα Απριλίου, η κανονική σεζόν ολοκληρώνεται. Κατά τη διάρκεια αυτής της περιόδου αρχίζει η ψηφοφορία για τα διάφορα μεμονωμένα αλλά και ομαδικά βραβεία.

1.2.3 Playoffs

Τα playoffs του NBA ξεκινούν μετά την ολοκλήρωση της κανονικής περιόδου με τις οκτώ (8) πρώτες ομάδες κάθε περιφέρειας, δηλαδή εκείνες με το καλύτερο ρεκόρ νικών-ηττών, ανεξάρτητα από την κατηγορία (division) στην οποία ανήκουν, να διεκδικούν τον τίτλο της πρωταθλήτριας.

Η πρόκριση στα playoffs και οι διασταυρώσεις των ομάδων στηρίζονται αυστηρά στο ρεκόρ της κανονικής περιόδου. Υψηλότερο ρεκόρ για μια ομάδα σημαίνει ότι αντιμετωπίζει μια φαινομενικά ασθενέστερη στον πρώτο γύρο. Ο πρώτος στην κατάταξη ξεκινά τα playoffs αντιμετωπίζοντας τον όγδοο, ο δεύτερος τον έβδομο κ.ο.κ.. Η ομάδα με το καλύτερο ρεκόρ, σε κάθε σειρά αγώνων, έχει πλεονέκτημα έδρας. Κάθε σειρά είναι «best-of-seven», δηλαδή κρίνεται στις 4 νίκες. Σε κάθε γύρο, η σειρά ακολουθεί ένα μοτίβο 2-2-1-1-1 όσον αφορά την έδρα διεξαγωγής των αγώνων. Με άλλα λόγια, μια ομάδα με πλεονέκτημα έδρας παίζει στο γήπεδό της το πρώτο, δεύτερο, πέμπτο και έβδομο παιχνίδι ενώ αγωνίζεται εκτός έδρας στο τρίτο, τέταρτο και έκτο παιχνίδι. Για την ιστορία, από το 1985 έως το 2013, οι τελικοί του NBA ακολούθησαν ένα μοτίβο 2-3-2 όσον αφορά την έδρα.

Τα playoffs ακολουθούν μια μορφή τουρνουά, καθώς η νικήτρια ομάδα κάθε σειράς παιχνιδιών προκρίνεται στον επόμενο γύρο, ενώ η αντίπαλη αποκλείεται. Οι πρωταθλήτριες ομάδες κάθε περιφέρειας διεκδικούν τον τίτλο της πρωταθλήτριας ομάδας του NBA.

Το πρωτάθλημα άρχισε να εφαρμόζει την τρέχουσα μορφή του, με τις οκτώ (8) κορυφαίες ομάδες κάθε περιφέρειας να προκρίνονται ανεξάρτητα από την υποδιαίρεσή τους στις divisions, την αγωνιστική περίοδο 2015-16. Τα προηγούμενα χρόνια, τα τρία πρώτα εισιτήρια από κάθε περιφέρεια στα playoffs πήγαιναν στους νικητές της εκάστοτε κατηγορία (division).

1.3 Ύψος Συμβολαίων, Φόρος Πολυτελείας και Πολιτική Συμβολαίων

Το salary cap είναι το μέγιστο όριο στο συνολικό χρηματικό ποσό που επιτρέπεται στις ομάδες του NBA να πληρώνουν τους παίκτες τους. Όπως και πολλά επαγγελματικά αθλητικά πρωταθλήματα, το NBA έχει ανώτατο όριο μισθών για τον έλεγχο του κόστους και της ισοτιμίας οφέλους, που ορίζεται από τη συλλογική σύμβαση διαπραγμάτευσης του πρωταθλήματος (CBA). Το όριο αυτό υπόκειται σε ένα πολύπλοκο σύστημα κανόνων και εξαιρέσεων και υπολογίζεται ως ποσοστό των εσόδων του πρωταθλήματος από την προηγούμενη περίοδο (season) σύμφωνα με το CBA BREAKDOWN[11].

Η πλειοψηφία των αμερικανικών πρωταθλημάτων (NFL, NHL, MLS) έχει αυστηρό ανώτατο όριο μισθών και απαγορεύει στις ομάδες να το υπερβαίνουν. Σε αντίθεση με αυτό, το ανώτατο όριο του NBA διαθέτει αρκετές σημαντικές εξαιρέσεις που επιτρέπουν στις ομάδες να το υπερβούν για να υπογράψουν παίκτες, ωστόσο, υπόκεινται σε μειωμένα προνόμια στην περίοδο της free agency. Οι ομάδες που υπερβαίνουν, επιπλέον, ένα καθορισμένο φορολογικό επίπεδο υπόκεινται στον φόρο πολυτελείας (φόρος για κάθε δολάριο που δαπανάται πάνω από το ανώτατο όριο φόρου πολυτελείας).

[11] https://cbabreakdown.com/salary-cap-overview

Ενώ οι περισσότερες ομάδες του NBA κατέχουν συμβόλαια που αποτιμώνται πάνω από το salary cap, λίγες ομάδες φτάνουν τα επίπεδα του luxury tax. Τα φορολογικά έσοδα κατανέμονται ομοιόμορφα μεταξύ των ομάδων που δεν πληρώνουν φόρους, επομένως, εκτός της οικονομικής ενίσχυσης που δέχονται, υπάρχει συχνά ένα κίνητρο πολλών εκατομμυρίων δολαρίων στους ιδιοκτήτες των ομάδων μην πληρώνουν φόρο πολυτελείας.

Σύμφωνα με το ισχύον σύστημα, ο φόρος πολυτελείας αξιολογείται σε διαφορετικά επίπεδα με βάση το ποσό που μια ομάδα υπερβαίνει το όριο του φόρου πολυτελείας (luxury tax).

Πίνακας 1.2: Επίπεδο φόρων (Tax levels) από την αγωνιστική περίοδο 2013–2014

Amount over tax threshold	Standard tax per excess dollar	Repeat offender tax per excess dollar
$5 million to $9,999,999	$1.75	$2.75
$10 million to $14,999,999	$2.50	$3.50
$15 million to $19,999,999	$3.25	$4.25
Over $20 million	$3.75 + $0.5 per $5 million	$4.75 + $0.5 per $5 million

Πηγή: Ιστοσελίδα SI[12] και Wikipedia[13]

Το καθεστώς δεν είναι σωρευτικό, καθώς κάθε επίπεδο φόρου ισχύει μόνο για ποσά που υπερβαίνουν το όριο του συγκεκριμένου επιπέδου. Για παράδειγμα, μια ομάδα που είναι 8 εκατομμύρια δολάρια πάνω από το φορολογικό όριο θα πληρώσει 1,50 δολάρια για καθένα από τα πρώτα 5 εκατομμύρια δολάρια πάνω από το φορολογικό όριο και 1,75 δολάρια ανά δολάριο για τα υπόλοιπα 3 εκατομμύρια δολάρια. Για τους κατ' εξακολούθηση παραβάτες ο φόρος διαφοροποιείται (Πίνακας 1.2).

Τέλος, όσον αφορά τα συμβόλαια που υπογράφουν οι ομάδες με τους παίκτες δεν συντάσσονται ελεύθερα αλλά σύμφωνα με ένα σύνολο κανονισμών και διαφόρων οικονομικών επιπέδων που καθορίζονται από τη συλλογική σύμβαση διαπραγμάτευσης του πρωταθλήματος, προκειμένου όλες οι ομάδες να έχουν την ευκαιρία να γίνουν ανταγωνιστικές. Υπάρχουν, μάλιστα, ειδικές κατηγορίες συμβολαίων στις οποίες υπάγονται ομάδες και παίκτες που πληρούν συγκεκριμένα

[12] https://www.si.com/nba/2013/07/10/nba-salary-cap-luxury-tax-figures-2013-14
[13] https://en.wikipedia.org/wiki/NBA_salary_cap

χαρακτηριστικά, όπως για παράδειγμα το supermax συμβόλαιο των βετεράνων παικτών που πληρούν μερικές προϋποθέσεις και σύμφωνα με το οποίο υλοποιήσιμη το αρχικό ποσό φτάνει έως και πάνω από 30%-35% του ανώτατου ορίου μισθών (salary cap).

1.4 NBA Draft[14]

Το NBA draft είναι ένα ετήσιο γεγονός στο οποίο συμμετέχουν και οι 30 ομάδες και επιλέγουν παίκτες που έχουν το δικαίωμα να παίξουν στο NBA. Τέτοιοι παίκτες είναι είτε ερασιτέχνες παίκτες που έχουν παίξει κολεγιακό μπάσκετ στις Η.Π.Α., είτε ερασιτέχνες παίκτες που έχουν τελειώσει τουλάχιστον ένα χρόνο το λύκειο, είτε μη Αμερικανοί παίκτες που μπορούν να είναι ήδη επαγγελματίες.

Οι γύροι του NBA draft είναι δύο. Στον δεύτερο γύρο, οι ομάδες επιλέγουν με σειρά αντίστροφη του ρεκόρ τους τη χρονιά που τελείωσε. Δηλαδή, η χειρότερη επιλέγει 31η και η καλύτερη 60η. Στον πρώτο γύρο, βεβαίως, υπάρχει η λοταρία (lottery) που καθορίζει σε ποια σειρά θα επιλέξει η κάθε ομάδα. Η λοταρία γίνεται κάθε χρόνο, πριν το draft, ανάμεσα στις 14 ομάδες οι οποίες τη σεζόν που μόλις τελείωσε έμειναν εκτός playoffs (ή τις ομάδες που έχουν τα δικαιώματα των επιλογών αυτών των ομάδων). Έχοντας τα χειρότερα ρεκόρ στο πρωτάθλημα, όσο πιο χαμηλά βρίσκονται στην κατάταξη τόσο πιο πιθανό είναι να επιλέξουν παίκτη στις πρώτες θέσεις. Οι υπόλοιπες 16 ομάδες επιλέγουν από το νούμερο 15 έως 30, με σειρά αντίστροφη του ρεκόρ που είχαν στην κανονική περίοδο.

Ουσιαστικά, στη γενική εικόνα του NBA draft, όσο πιο «νωρίς» επιλέγει μια ομάδα παίκτη τόσο πιο χαμηλά στην κατάταξη βρέθηκε την ακριβώς περασμένη αγωνιστική περίοδο. Με αυτό τον τρόπο, οι ομάδες που δεν είχαν μια καλή αγωνιστική χρονιά έχουν την ευκαιρία να επιλέξουν στις πρώτες θέσεις έναν πολύ

[14] Πηγές: α) http://www.ageofbasketball.net/archives/3373,
β) https://www.sportime.gr/basket/nba/nba-draft-i-lotaria-etsi-litourgi/,
γ) https://en.wikipedia.org/wiki/National_Basketball_Association.

καλό παίκτη, προκειμένου να ενισχυθεί και να έχει βελτιωθεί αγωνιστικά. Υπάρχουν ομάδες που επιλέγουν να υποστηρίξουν το draft pick τους και να «χτίσουν» μια ομάδα πάνω σε αυτό μελλοντικά, ενώ άλλες εκμεταλλεύονται τα draft picks ανταλλάσσοντάς τα προκειμένου να γίνουν ανταγωνιστικές άμεσα.

Τα συμβόλαια των επιλογών του πρώτου γύρου draft είναι εγγυημένα για δυο χρόνια, ενώ η ομάδα έχει τη δυνατότητα να αποφασίσει για το αν θα συνεχίσει τη συνεργασία με τον συγκεκριμένο παίκτη για την τρίτη και τέταρτη σεζόν. Το ύψος των συμβολαίων αυτών είναι προκαθορισμένο και κλιμακώνεται ανάλογα με τη θέση, δηλαδή η πρώτη θέση αντιστοιχεί στο υψηλότερο rookie συμβόλαιο και η τριακοστή στο χαμηλότερο. Από την άλλη, τα συμβόλαια των επιλογών του δεύτερου γύρου στο draft δεν είναι καθορισμένα.

1.5 Ανταλλαγές (Trades)

Κάθε αγωνιστική περίοδο οι ομάδες μπορούν να ανταλλάσσουν ή να πωλούν μεταξύ τους συμβόλαια παικτών, δικαιώματα παικτών, δικαιώματα draft ή draft picks. Υπάρχουν κανονισμοί του πρωταθλήματος που καθιστούν ισάξιες και δίκαιες ανταλλαγές. Οι κανονισμοί αυτοί καθορίζονται από τη συλλογική σύμβαση διαπραγμάτευσης του πρωταθλήματος (CBA)[15]. Η πολιτική των ομάδων όσον αφορά τις ανταλλαγές (trade policy) είναι ιδιαίτερα σημαντική και εξαρτάται ακόμα και από τις αποφάσεις που έχουν ληφθεί σε προηγούμενες σεζόν.

1.6 Αντικείμενο, στόχοι και δομή της εργασίας

Το αντικείμενο της παρούσας διπλωματικής εργασίας είναι η μελέτη των παραγόντων που επιδρούν στο πλήθος νικών που πετυχαίνει μια ομάδα

[15]Βλ. https://cbabreakdown.com/trades

καλαθοσφαίρισης του NBA από τους 82 αγώνες της κανονικής περιόδου του πρωταθλήματος. Στόχος είναι η εκτίμηση ενός υποδείγματος με εξαρτημένη μεταβλητή το πλήθος νικών, στο οποίο λαμβάνονται υπόψη παράγοντες που σχετίζονται με τις διοικητικές και οικονομικές αποφάσεις των ομάδων εκτός από τους αγωνιστικούς. Επίσης, η εργασία αποσκοπεί στη μελέτη της επίδρασης της αγωνιστικής εικόνας μιας ομάδας από αυτή που είχε την προηγούμενη χρονιά η ίδια η ομάδα αλλά και οι αντίπαλοί της.

Για τον σκοπό αυτό εκτιμάται ένα Δυναμικό Υπόδειγμα Durbin Χωρικής Διάχυσης με τη Μέθοδο Μεγίστης Πιθανοφάνειας (ML) στο πρόγραμμα Stata. Η Μεθόδος Μεγίστης Πιθανοφάνειας επιλέχθηκε για την αποφυγή του προβλήματος ενδογένειας, κάτι που στην έως τώρα βιβλιογραφία δεν έχει ληφθεί υπόψη.

Η βάση δεδομένων αποτελείται από τις 30 ομάδες που συμμετέχουν στο πρωτάθλημα της Αμερικής, ενώ επιλέχθηκε μία χρονοσειρά δεκαπέντε ετών, από την αγωνιστική περίοδο 2005-2006 έως 2019-2020, καθώς σε αυτό το διάστημα αγωνίζονται και οι 30 ομάδες.

Η δομή της εργασίας είναι ως εξής:

- Το δεύτερο κεφάλαιο αφορά τη βιβλιογραφική ανασκόπηση τόσο για την αθλητική στατιστική όσο και για τη μεθοδολογία που εφαρμόζεται στην παρούσα μελέτη.
- Το τρίτο κεφάλαιο αφορά το μεθοδολογικό πλαίσιο, όπου περιγράφονται το βασικό υπόδειγμα μελέτης, η μέθοδος εκτίμησής του καθώς και οι απαραίτητοι έλεγχοι.
- Το τέταρτο κεφάλαιο αφορά τα εμπειρικά αποτελέσματα, όπου γίνεται η εφαρμογή του υποδείγματος και η ανάλυση των αποτελεσμάτων.
- Και, τέλος, στο πέμπτο κεφάλαιο παρουσιάζεται μια ανακεφαλαίωση των συμπερασμάτων από τα εμπειρικά ευρήματα του προηγούμενου κεφαλαίου.

Κεφάλαιο 2: Βιβλιογραφική Ανασκόπηση

2.1 Αθλητική Στατιστική

Κατά καιρούς έχουν πραγματοποιηθεί μελέτες αναφορικά με την αθλητική στατιστική και την ανάλυση δεδομένων που λαμβάνουμε είτε από ατομικά είτε από ομαδικά αθλήματα. Ίσως η πιο διάσημη μεταξύ αυτών, λόγω της μεταφοράς της στη μεγάλη οθόνη, είναι το βιβλίο «Moneyball» του Michael Lewis (2003) που αναλύει την επιτυχία των Oakland Athletics στο MLB (Major League Baseball), μια ομάδα που με τη χρήση ερμηνευτικών, αγωνιστικών δεδομένων και στατιστικών έγινε ανταγωνιστική παρά το μικρό της budget.

Όσον αφορά το μπάσκετ, ο άνθρωπος που καθιέρωσε τη σημασία και χρήση των στατιστικών είναι ο Dean Oliver, ο οποίος το 2004 επεκτάθηκε στη φιλοσοφία των "Four Factors" από το βιβλίο του 2002 «Basketball on Paper», σε μια προσπάθεια να προσδιορίσει πώς τέσσερις σημαντικές στρατηγικές σχετίζονται με την επιτυχία στο μπάσκετ. Στην ουσία πρόκειται για οκτώ παράγοντες, τέσσερις επιθετικούς και τους αντίστοιχους τέσσερις αμυντικούς. Ο Dean Oliver καθόρισε ένα συγκεκριμένο σύνολο βαρών σε καθεμία από τις τέσσερις κατηγορίες για να εξακριβώσει την αξία κάθε παράγοντα σε σχέση με τη νίκη. Οι παράγοντες αυτοί κατά σειρά φθίνουσας σημαντικότητας ως προς το αποτέλεσμα του αγώνα αφορούν το σουτ, τα λάθη, τα rebounds, τις ελεύθερες βολές και είναι οι εξής: αποτελεσματικότητα εντός πεδιάς (effective field goal percentage-eFG%), λάθη ανά κατοχή (turnovers per possession-TPP), ποσοστό επιθετικών ριμπάουντ της ομάδας από των συνολικών επιθετικών ριμπαόυντ (offensive rebounding percentage-ORP) και πλήθος επιθέσεων που καταλήγουν σε βολές (free throw rate-FTR). Ο Dean Oliver αναφέρεται αρκετές φορές ως ο «πατέρας» των στατιστικών στον κόσμο του μπάσκετ. Αυτό, μάλιστα, επιβεβαιώνεται από τις έρευνες που ακολούθησαν έπειτα από τη δημοσίευση του για τους «Four Factors» (2004) που αποτέλεσαν αφορμή για πλήθος μελετών.

Μία, ακόμα, σημαντική προσωπικότητα για την ανάπτυξη της στατιστικής στο μπάσκετ είναι ο John Hollinger (Pro Basketball 2002, 2003, 2004, 2005) ο οποίος στα κείμενά του αφιερώθηκε σε μία εκτενή ανάλυση και παρουσίαση στατιστικών που καταγράφονται στους αγώνες μπάσκετ αλλά με μια πιο ουσιαστική ματιά, αποσκοπώντας στο σχηματισμό μιας συνολικής εικόνας για τους παίκτες και τις ομάδες. Βασιζόμενος στον παράγοντα των κατοχών, ο Hollinger, πετυχαίνει σε σημαντικό βαθμό την εμβάθυνση στην αμυντική και επιθετική ποιότητα των παικτών σε ατομικό επίπεδο αλλά και της εκάστοτε ομάδας στο σύνολό της. Επίσης, πραγματοποιεί προβλέψεις βάσει στατιστικών, σε αγωνιστικό επίπεδο. Για παράδειγμα, το πώς μια ομάδα θα πρέπει να θεμελιώσει την επίθεσή της, προκειμένου να είναι αποδοτική, πάντα με στόχο τη νίκη. Να σημειωθεί, τέλος, ότι ο Hollinger είναι γνωστός για τη δημιουργία του στατιστικού δείκτη Player Efficiency Rating (PER). Στην προσπάθειά του να συμπεριλάβει όλες τις συνεισφορές ενός παίκτη σε έναν αριθμό, πρότεινε τη συγκεκριμένη λεπτομερή φόρμουλα που αξιολογεί τη στατιστική απόδοση κάθε παίκτη και η οποία χρησιμοποιείται πλέον ευρέως σε όλες τις διοργανώσεις.

Ωστόσο, η αφετηρία της αθλητικής στατιστικής είχε τεθεί νωρίτερα. Στο χώρο του αθλητισμού ιδιαίτερο ενδιαφέρον έχει τόσο η πρόβλεψη του νικητή όσο και του τελικού σκορ. Αυτό συνδέεται άμεσα με το ζήτημα μοντελοποίησης της δυναμικότητας κάθε παίκτη ή κάθε ομάδας για τα ομαδικά αθλήματα. Ο Elo (1978) μοντελοποίησε την αγωνιστική «εμφάνιση» (R_p) ενός παίκτη σκάκι με ένα απλό γραμμικό μοντέλο, όπου ο δείκτης R_p της εμφάνισης ακολουθεί κανονική κατανομή. Οι Glickman (1993) και Fahrmeir and Tutz (1994) εισήγαγαν τα δυναμικά μοντέλα σύγκρισης σε ζεύγη που εφαρμόστηκαν σε πρώιμο στάδιο στο soccer από τους Fahrmeir and Tutz (1994) και Koopman and Lit (2015), στο σκάκι και στο τένις από τον Glickman (1999), στο ποδόσφαιρο από τους Glickman and Stern (1998) και Glickman (2001) και στο μπάσκετ από τους Knorr-Held (2000) υπολογίζοντας μία ποσότητα που εκφράζει τη δυναμική της ομάδας ή των παικτών βάσει της μεταβλητότητας του χρόνου.

Έχοντας, πλέον, τεθεί μερικά θεμέλια της αθλητικής στατιστικής, ακολούθησαν αρκετές μελέτες που επικεντρώθηκαν στον βαθμό σημαντικότητας κάποιων παραγόντων-μεταβλητών για τη νίκη μιας ομάδας, ενώ δεν είναι λίγες οι έρευνες που πραγματεύονται την πρόβλεψη του αποτελέσματος ενός αγώνα.

Εστιάζοντας στο μπάσκετ, η Stefani (1977a, b) εισήγαγε την προσέγγιση που βασίζεται στην παλινδρόμηση με τη Μέθοδο Ελαχίστων Τετραγώνων με την οποία, αφού αρχικά πρότεινε έναν τρόπο αξιολόγησης των ομάδων του κολεγιακού πρωταθλήματος καλαθοσφαίρισης αλλά και ποδοσφαίρου χρησιμοποιώντας το μέσο όρο των αξιολογήσεων του αντιπάλου με περιθώριο νίκης, ύστερα προχώρησε σε προβλέψεις για τα αποτελέσματα των αγώνων.
Το κολεγιακό πρωτάθλημα (NCAA) αποτέλεσε αντικείμενο μελέτης για αρκετούς μελετητές, όπως οι Schwertman et al. (1991), ο Carlin (1991) και ο Harville (2003), οι οποίοι με την εφαρμογή γραμμικών και λογιστικών παλινδρομήσεων, εκτίμησαν τις ακριβείς βαθμολογίες των ομάδων, καθώς και της πιθανότητας νίκης τους, αντίστοιχα. Οι Schwertman et al. (1991) υπολόγισαν τις πιθανότητες με βάση την υπόθεση ότι ο παράγοντας της δυναμικότητας των ομάδων κατανέμεται κανονικά, ενώ ο Carlin (1991) βελτιώνει μοντέλα πιθανότητας που προτάθηκαν, όπως των Schwertman et al. (1991), εκμεταλλευόμενος εξωτερικές πληροφορίες σχετικά με τη σχετική δυναμικότητα των ομάδων. Ο Harville (2003) προτείνει την εφαρμογή ελαχίστων τετραγώνων σε ένα στατιστικό μοντέλο στο οποίο η αναμενόμενη διαφορά στο σκορ σε κάθε παιχνίδι διαμορφώνεται ως διαφορά στις ομαδικές επιδράσεις λαμβάνοντας υπόψη την έδρα του αγώνα. Μάλιστα, με ορισμένες τροποποιήσεις που εφάρμοσε το υπόδειγμα πρόβλεψης που πρότεινε μπορεί να ανταγωνιστεί μοντέλα πρόβλεψης στοιχηματικών εταιρειών. Ο Stern (1994) προτείνει μία μέθοδο πρόβλεψης του αποτελέσματος ενός παιχνιδιού, δεδομένου του σκορ και του υπολειπόμενου χρόνου μέχρι το τέλος του παιχνιδιού, βασισμένη στην κίνηση Brown, εισάγοντας μια νέα μεθοδολογία από αυτήν που «επικρατούσε» έως τότε στο συγκεκριμένο αντικείμενο μελέτης.

Σε ορισμένες έρευνες, χωρίς την εφαρμογή κάποιας διαφορετικής και εναλλακτικής μεθοδολογίας από αυτές που έχουν αναφερθεί ήδη, λήφθηκε υπόψη και ο

παράγοντας της έδρας ως πλεονέκτημα για τις γηπεδούχους ομάδες έναντι των φιλοξενούμενων. Παράδειγμα τέτοιων μελετών αποτελούν αυτές των Harville and Smith (1994), Jones (2007, 2008) και Entine and Small (2008).

Ο δείκτης των Πυθαγορείων νικών (Bill James, 1980) συνιστά μια εναλλακτική μέθοδο πρόβλεψης του ποσοστού των νικών μιας ομάδας. Πρόκειται για εφεύρεση του Bill James στο μπέιζμπολ, ο οποίος στηρίχτηκε στο γεγονός ότι τα ποσοστά νίκης της ομάδας είναι γενικά στενά συνδεδεμένα με τους πόντους που πετυχαίνει και τους πόντους που δέχεται. Το Πυθαγόρειο ποσοστό νίκης (Pyth) υπολογίζεται ως εξής:

$$PYTH_t = \frac{PTS_t^X}{PTS_t^X + PTS_o^X}$$

όπου ο δείκτης *t* εξακολουθεί να υποδεικνύει ομάδα, ο δείκτης *o* υποδεικνύει τον αντίπαλο και ο δείκτης *x* είναι ένας εκθέτης που προσδιορίζεται εμπειρικά. Στο μπέιζμπολ, ο James (1985) κατέληξε εμπειρικά ότι ο εκθέτης *x* = 2, οδηγώντας στη χρήση του δείκτη.

Σε μεταγενέστερες εργασίες διαπιστώθηκε ότι ένας ελαφρώς χαμηλότερος εκθέτης περίπου 1,8 λειτουργεί καλύτερα. Στο NBA η τιμή του *x* έχει εκτιμηθεί εμπειρικά από διαφορετικές μελέτες να είναι μεταξύ 13 και 17 περίπου (π.χ. Oliver, 1996, CoolStandings, 2006, για παράδειγμα). Η τιμή διαφέρει ανάλογα με την περίοδο. Συγκεκριμένα, τα περιθώρια νίκης έχουν αλλάξει λίγο τα τελευταία χρόνια. Αυτό έχει ως αποτέλεσμα να είναι απαραίτητοι μικρότεροι εκθέτες για να συσχετιστούν οι πόντοι με το ποσοστό νικών. Η επίδραση αυτή, όμως, αν ληφθούν υπόψη οι πολύ καλές και πολύ κακές ομάδες από τις οποίες υπάρχουν μόνο μερικές κάθε σεζόν, γίνεται πολύ καλύτερη με έναν μεγαλύτερο εκθέτη. Οι τυπικές εκτιμήσεις σφάλματος του εκθέτη με τη μέθοδο ελαχίστων τετραγώνων τείνουν να αποδίδουν μικρότερο εκθέτη επειδή οι νίκες των περισσότερων ομάδων συσσωρεύονται μεταξύ περίπου 30 και 50 σε πλήθος (που αντιστοιχούν σε ποσοστά μεταξύ περίπου 38% και 62%). Ωστόσο, αυξάνοντας τον εκθέτη δεν θέτει σε κίνδυνο σημαντικά την πρόβλεψη των νικών των πολύ καλών ή πολύ κακών ομάδων και, ως εκ τούτου, ο Oliver (ESPN 2007) καθιέρωσε τη χρήση του εκθέτη 16.5.

Σε άλλα πρωταθλήματα, συμπεριλαμβανομένου του κολλεγιακού, του σχολικού και του WNBA, μικρότεροι εκθέτες έχει βρεθεί ότι λειτουργούν καλύτερα, καθώς είναι λιγότερες οι κατοχές σε ένα τυπικό παιχνίδι σε αυτά τα πρωταθλήματα (βλέπε Pomeroy (2006), για παράδειγμα). Επίσης, να σημειωθεί ότι λόγω της ισότητας των κατοχών για μια ομάδα και των αντιπάλων της, το PTS_t και το PTS_o μπορούν να αντικατασταθούν στην παραπάνω εξίσωση από τα Offensive Rating (*Ortg*) και Defensive Rating (*Drtg*).

Συγκριτικά με τις «Πυθαγόρειες νίκες», η μέθοδος Bell Curve (που εισήχθη από τον Dean Oliver, 2004) είναι μια πιο θεωρητική προσέγγιση των πόντων που σημείωσε μια ομάδα και αυτών που δέχτηκε από τον αντίπαλο σχετικά με το ποσοστό νικών της. Σε αντίθεση με τη μέθοδο των Πυθαγόρειων νικών, η συγκεκριμένη μέθοδος βασίζεται στην υπόθεση ότι οι πόντοι που πετυχαίνει μια ομάδα και αυτοί που δέχεται είναι κανονικά κατανεμημένοι, και αφαιρώντας τους μπορούν να σχηματίσουν μία άλλη κανονικά κατανεμημένη τυχαία μεταβλητή, που εκφράζουν τους «καθαρούς πόντους».[16]

Ο τύπος για την πρόβλεψη του ποσοστού νίκης για την ομάδα *t* (Win_t *%*) της μεθόδου αυτής:

$$Win\% = NORMDIST\left[\frac{PPG_t - PPG_o}{StDev(PPG_t - PPG_o)}\right]$$

όπου το PPG_t είναι οι πόντοι ανά παιχνίδι για την ομάδα *t*, το PPG_o είναι οι πόντοι ανά παιχνίδι για τους αντιπάλους ο, StDev ($PPG_t - PPG_o$) είναι η τυπική απόκλιση των καθαρών πόντων ($PPG_t - PPG_o$) σε όλη την αγωνιστική season μιας ομάδας, και το *NORMSDIST* είναι η κανονική αθροιστική κατανομή συνάρτηση και αντιπροσωπεύει την περιοχή κάτω από την τυπική κανονική "καμπύλη καμπάνας" προς τα αριστερά της αξίας σε παρένθεση. Τα πλεονεκτήματα της μεθόδου αυτής σε σχέση με τις «Πυθαγόρειες νίκες» είναι ότι δεν χρειάζεται εμπειρική τροποποίηση για εφαρμογή σε διαφορετικά πρωταθλήματα ή διαφορετικές χρονικές και αγωνιστικές περιόδους ενώ, παράλληλα, ενσωματώνει πληροφορίες

[16] Οι «καθαροί πόντοι» κανονικοποιούνται διαιρώντας με την τυπική απόκλιση των «καθαρών πόντων» για τη συγκεκριμένη ομάδα, σχηματίζοντας έτσι μια Z-στατιστική. Η εκτιμώμενη πιθανότητα νίκης δίνεται στη συνέχεια από την πιθανότητα μια τυχαία μεταβλητή που ακολουθεί την τυπική κανονική κατανομή να λαμβάνει μια τιμή λιγότερη από αυτή τη Z-στατιστική.

σχετικά με το πόσο περισσότερο ή λιγότερο σκοράρουν οι ομάδες έναντι των αντιπάλων τους.

Οι Berri et al. (2006) εξετάζουν τον αθλητισμό με τη χρήση στατιστικών εργαλείων και μεθόδων και προσπαθούν να διερευνήσουν τους σημαντικούς παράγοντες για τη νίκη των παιχνιδιών από στατιστική άποψη, ενώ, συγχρόνως, καταπιάνονται με τη δημιουργία στατιστικών δεικτών. Οι Berri et al. (2006) καθιστούν απολύτως σαφές ότι μεγάλο μέρος της λήψης αποφάσεων των προπονητών και των γενικών διευθυντών δεν βασίζεται στην ανάλυση των αριθμών. Οι συγγραφείς υποθέτουν ότι οι παίκτες δεν επηρεάζονται αγωνιστικά από τους συμπαίκτες τους, γεγονός που τους επιτρέπει να αναπτύξουν μια γραμμική διατύπωση βαρών καθορίζοντας πρώτα τη σχέση μεταξύ νικών και ομαδικών στατιστικών στοιχείων και στη συνέχεια υποθέτοντας μια παρόμοια σχέση μεταξύ νικών και ατομικών στατιστικών των παικτών.

Υπάρχουν εργασίες που βασίστηκαν στους «Four Factors» του Dean Oliver (2004), όπως του Küpfer (2005), στις οποίες τοποθετήθηκαν διάφορα βάρη σε κάθε έναν από τους τέσσερις προαναφερθέντες παράγοντες, προκειμένου να εξεταστεί η σημασία αυτών σε σχέση με τις νίκες που πετυχαίνει μία ομάδα. Αυτή η μελέτη, όπως και άλλες αντίστοιχες, απλά επιβεβαίωσαν τα αποτελέσματα του Dean Oliver ως προς τη σημαντικότητα των τεσσάρων παραγόντων.

Οι Kubatko et al. (2007) συγκέντρωσαν τις βασικές μεταβλητές της ανάλυσης μπάσκετ που απορρέουν από τα βασικά στατιστικά και καθόρισαν ένα πλαίσιο στο οποίο μπορούν να θεμελιωθούν επόμενες έρευνες. Πολύ σημαντικοί είναι, φυσικά, οι «Four Factors» του Dean Oliver (2004) για τους οποίους οι Kubatko et al. (2007) παρατήρησαν ότι η μέση τιμή τους διαφέρει με την πάροδο του χρόνου, και μάλιστα παρατήρησαν μια γενική συνολική μείωση του αναμενόμενων επιθετικών δεικτών και προφανώς των λαθών (turnovers). Επίσης, αξιοσημείωτη είναι η εισαγωγή των προηγμένων στατιστικών (advanced statistics)[17] που υπολογίζονται μέσω box-score δεδομένων και τα οποία έχουν πλέον καθιερωθεί στον

[17] βλ. ΠΑΡΑΡΤΗΜΑ ΙΙΙ: Προηγμένα Στατιστικά (Advanced Stats)

«στατιστικό» κόσμο του μπάσκετ. Σε αυτό το σημείο να υπογραμμιστεί ότι οι Kubatko et al. (2007) βασίστηκαν στο στατιστικό δεδομένο των κατοχών (possessions) ακολουθώντας τους προκατόχους τους Dean Oliver (2004) και Hollinger (2002, 2003, 2004, 2005).

Καθώς με την πολυμεταβλητή ανάλυση καθίσταται εφικτή η βελτίωση της μελέτης της επίδρασης των τεσσάρων παραγόντων και όχι μόνο στη νίκης μιας ομάδας, οι Doolittle και Pelton (2010) ερεύνησαν τα στατιστικά στοιχεία της ομάδας για αυτούς τους παράγοντες και προσπάθησαν να τους συσχετίσουν με τη νίκη, ή ακόμα και να αναπτύξουν μία μέθοδο πρόβλεψης του αποτελέσματος βασισμένη σε μεμονωμένες, αγωνιστικές προβλέψεις για κάθε παίκτη.

Οι Teramoto και Cross (2010) χρησιμοποιώντας, επίσης, πολυμεταβλητή ανάλυση, εκτίμησαν την επίδραση των επιθετικών και αμυντικών τεσσάρων παραγόντων στο ποσοστό νικών μιας ομάδας μπάσκετ του NBA στην κανονική περίοδο αλλά και στα playoffs. Γενικά, επιβεβαίωσαν τα αποτελέσματα του Küpfer (2005) για την επίδραση των τεσσάρων παραγόντων. Διαπιστώθηκε, λοιπόν, ότι το επιθετικό και αμυντικό Effective Field Goals (*eFG*) ήταν οι περισσότερο στατιστικά σημαντικοί παράγοντες για τη νίκη τόσο στην κανονική περίοδο όσο και στα playoffs, με το αμυντικό *eFG* πιο σημαντικό για τα playoffs. Από την άλλη, το επιθετικό και αμυντικό Free Throw Rate (*FTR*), παρά τη σημαντικότητά τους για την πρόβλεψη του αποτελέσματος νίκης ή ήττας, ήταν σχετικά το λιγότερο σημαντικό από τους τέσσερις παράγοντες.

Γενικά, έχουν πραγματοποιηθεί ποικίλες μελέτες για την ανάλυση των παραγόντων που επιδρούν στις νίκες μιας ομάδας μπάσκετ αλλά και την πρόβλεψη του αποτελέσματος ενός αγώνα, οι οποίες μεθοδολογικά στηρίζονται, κατά κύριο λόγο, στη Μέθοδο Ελαχίστων Τετραγώνων. Δεν αφορούν μόνο στο NBA αλλά και σε άλλες διοργανώσεις ενώ παρουσιάζουν διαφοροποιημένο σε μικρό βαθμό το αντικείμενο έρευνας. Ένα παράδειγμα τέτοιας μελέτης στην ισπανική λίγκα αποτελεί αυτή των Javier García et al. (2013), οι οποίοι προσπάθησαν να προσδιορίσουν τους δείκτες που διακρίνουν μια νικήτρια ομάδα από την ηττημένη στην κανονική περίοδο, έναντι των playoffs.

Ο Kline (2005) ξεφεύγοντας τόσο από την ανάλυση των ελαχίστων τετραγώνων όσο και από τους αγωνιστικούς παράγοντες που λαμβάνονται υπόψη από την πλειοψηφία των ερευνών, εισήγαγε στη μελέτη κριτηρίων που οδηγούν μια ομάδα σε νίκη τον οικονομικό παράγοντα του salary. Εφαρμόζοντας μοντέλα δομικών εξισώσεων (structural equation modeling, SEM) προσπάθησε να εκτιμήσει το «πραγματικό», ουσιαστικό επίπεδο της «δομής μιας κατασκευής», όπως για παράδειγμα στην προκειμένη περίπτωση την επιθετική ή αμυντική ποιότητα μιας ομάδας, με βάση διάφορους δείκτες. Εκτιμώνται, λοιπόν, όπως σε ένα σύνηθες μοντέλο παλινδρόμησης, ποικίλοι παράγοντες ενσωματώνοντας όμως αναλύσεις path analysis μεταξύ των μεταβλητών και των εκτιμώμενων αυτών παραγόντων. Έτσι, επιτυγχάνεται η συσχέτιση των παραγόντων της μελέτης με το μισθό (salary) και έπειτα με τις νίκες μιας ομάδας.

Λίγα χρόνια αργότερα, ο Tarek Al Baghal (2012), στηριζόμενος σε προηγούμενες μελέτες και επηρεασμένος από τον Kline (2005), προσπάθησε, αρχικά, να εισάγει και να καθιερώσει τα μοντέλα δομικών εξισώσεων (SEM) ως αποτελεσματικό στατιστικό εργαλείο στον τομέα των sport analytics and statistics. Ασχολήθηκε με την ανάλυση παλινδρόμησης για τη νίκη και τους τέσσερις παράγοντες χρησιμοποιώντας ένα μεγαλύτερο αριθμό αγωνιστικών περιόδων από τους Teramoto και Cross (2010) ενώ, ακόμα, χρησιμοποίησε μοντέλα δομικών εξισώσεων SEM, για να καθορίσει εάν οι τέσσερις παράγοντες θα μπορούσαν να μοντελοποιηθούν ως δείκτες πιο γενικών μεμονωμένων μεταβλητών που αφορούν στην επιθετική και αμυντική ποιότητα των ομάδων. Ακολούθως, συσχέτισε τις δύο γενικές μεταβλητές- χαρακτηριστικά μιας ομάδας με τις νίκες της. Επιπρόσθετα, οι μισθοί της ομάδας ενσωματώθηκαν στα μοντέλα SEM, προκειμένου να εκτιμηθεί η σχέση κόστους με την επιθετική και αμυντική ποιότητα καθώς και με το ποσοστό νίκης. Ο Baghal (2012) διαπίστωσε ότι οι συνολικοί μισθοί κάθε ομάδας έχουν υψηλή επίδραση στην επιθετική της ποιότητα, ενώ επιδρά σε μέτριο βαθμό στο ποσοστό νικών της ομάδας. Συμπέρανε, λοιπόν, ότι οι συνολικοί μισθοί έχουν έμμεση επίδραση στις νίκες της ομάδας καθώς υψηλοί μισθοί συνεπάγονται

υψηλού επιπέδου επιθετική ποιότητα και με τη σειρά της συνεπάγεται υψηλό ποσοστό νικών.

Οι Lori Hoffman and Maria Joseph (2003) χρησιμοποίησαν την Ανάλυση Κύριων Συνιστωσών (Principal Component Analysis) για να προσδιορίσουν τους πιο σημαντικούς παράγοντες που θα οδηγήσουν μια ομάδα να προκριθεί στα playoffs. Επιπρόσθετα, προέβλεψαν την πιθανότητα μιας ομάδας να τερματίσει σε μια από τις θέσεις που οδηγούν στην post season. Οι παράγοντες που συμπεριέλαβαν στην έρευνά τους είναι τόσο αγωνιστικοί όσο και μη αγωνιστικοί: Μέσοι πόντοι που πέτυχε μια ομάδα ανά παιχνίδι (Average points scored per game), Μέσοι πόντοι που δέχτηκε μια ομάδα ανά παιχνίδι (Average points allowed per game), Ποσοστό εύστοχων εντός πεδιάς (Team percentage of field goals made), Πλήθος ετών που η ομάδα είναι ενεργή (Number of years since team's establishment), Αξιολόγηση των συνολικών ετήσιων μισθών των ομάδων (Rank of team's total yearly payroll), Ρεκόρ νίκης-ήττας στο NBA του πρώτου προπονητή της ομάδας (Head Coach's NBA record), Διαφορά αμυντικών και επιθετικών λαθών (Defensive turnovers less offensive turnovers), Ποσοστό νικών της ομάδας την περασμένη αγωνιστική περίοδο (Last season's percentage of games won), Μέσος αριθμός θεατών ανά αγώνα (Average attendance per game), Πλήθος ριμπάουντ ανά αγώνα (Number of rebounds per game), Ποσοστό νέων παικτών στο ρόστερ (Ratio of new players), Διάμεσος της ηλικίας των παικτών της ομάδας (Median age of team).

Συνεχίζοντας, ακολουθούν δύο εργασίες με χρήση νευρωνικών δικτύων για την μελέτη των παραγόντων που σχετίζονται με τις νίκες. Αρχικά, ο Renato Amorim Torres (2013) προσπάθησε να προβλέψει τα αποτελέσματα του επερχόμενου παιχνιδιού με τη χρήση γραμμικής παλινδρόμησης και, συγκεκριμένα, με τη Μέθοδο Μεγίστης Πιθανοφάνειας καθώς και μέσω ενός Νευρωνικού Δικτύου Πολλαπλών Στρωμάτων, τα αποτελέσματα των οποίων και συνέκρινε. Στο πλαίσιο αυτό, ο Jaak Uudmae (2017) προσπάθησε να προβλέψει τα αποτελέσματα του επερχόμενου παιχνιδιού (νίκη ή ήττα). Συνέκρινε τα αποτελέσματα διαφορετικών μεθόδων, συμπεριλαμβανομένης των μηχανών διανυσμάτων υποστήριξης (support vector machine-SVM), της γραμμικής παλινδρόμησης και παλινδρόμηση νευρωνικού δικτύου (neural network regression-NNR).

Στις προηγούμενες τρεις μελέτες των Lori Hoffman and Maria Joseph (2003), Renato Amorim Torres (2013) και Jaak Uudmae (2017) όπου έγινε μια προσπάθεια μεθοδολογικής καινοτομίας, παρατηρείται ότι είναι περισσότερο προσανατολισμένες στα αποτελέσματα και δεν δίνεται μεγάλη προσοχή στη στατιστική δομή των δεδομένων. Ειδικότερα, στην εφαρμογή του γραμμικού μοντέλου, θεωρήθηκε ότι τα δεδομένα είναι ανεξάρτητα γεγονός που δεν αληθεύει δεδομένου ότι οι ομάδες συμμετέχουν σε διαφορετικά παιχνίδια (παρατηρήσεις). Έτσι, ο Xiao Zhang (2019), βασιζόμενος στην υπόθεση ύπαρξης συσχέτισης μεταξύ των παρατηρήσεων, εισάγει ένα στατιστικό μοντέλο με τη μέθοδο Γενικευμένων Εξισώσεων Εκτίμησης (Generalized Estimating Equation-GEE) με στόχο την εκτίμηση των αποτελεσμάτων των παιχνιδιών και την επεξήγηση των περισσότερο σημαντικών παραγόντων.

Αξιοσημείωτες είναι δύο μελέτες που καινοτόμησαν ως προς τα δεδομένα που χρησιμοποίησαν. Κατ' αρχάς, η έρευνα του Yuanhao (Stanley) Yang (2015), στην οποία συμπεριλήφθηκαν και τα ατομικά στατιστικά των παικτών με μεγαλύτερη βαρύτητα. Συγκεκριμένα, ο Yang (2015) εφάρμοσε πολλαπλή γραμμική παλινδρόμηση ελαχίστων τετραγώνων για να συσχετίσει, αρχικά, τα ατομικά στατιστικά των παικτών, και συγκεκριμένα τον δείκτη PER που μετράει την αποδοτικότητα του παίκτη, με το αποτέλεσμα του παιχνιδιού για την ομάδα του (νίκη ή ήττα). Στη συνέχεια, ανέπτυξε ένα μοντέλο πρόβλεψης αποτελεσμάτων της κανονικής περιόδου. Ένα χρόνο μετά, ο Eric Scot Jones (2016) προχώρησε σε μελέτη με χρήση in-game στατιστικών, σε αντίθεση με όλες τις έρευνες που προηγήθηκαν. Ασχολήθηκε με την εκτίμηση ενός μοντέλου που εξηγεί το final point spread του παιχνιδιού μέσω των in-game στατιστικών και ενός μοντέλου πρόβλεψης της νίκης ή ήττας αναλύοντας τους παράγοντες in-game που οδηγούν στη νίκη.

Τέλος, μία εργασία που σχετίζεται, σε κάποιο βαθμό, με την έρευνα που πραγματοποιήθηκε και παρουσιάζεται στη συγκεκριμένη εργασία είναι του Hans Manner (2015). Βασιζόμενος στις γραμμικές μοντελοποιήσεις που είχαν αναπτυχθεί έως το 2015, ο Manner (2015) εισήγαγε τον παράγοντα της έδρας του παιχνιδιού

αλλά και των back-to-back αγώνων. Το μοντέλο που προτείνει εκτιμά τη δυναμικότητα της ομάδας συναρτήσει της μεταβλητότητας του χρόνου για το διάστημα οκτώ αγωνιστικών περιόδων (δηλαδή ετών). Πρόκειται για μια δυναμική Γκαουσιανή αυτοπαλίνδρομη διαδικασία. Ο Manner (2015) συμπέρανε ότι η δυναμικότητα μιας ομάδας διατηρείται στη διάρκεια του συνόλου των αγωνιστικών περιόδων και όχι σε κάθε μεμονωμένη σεζόν. Επίσης, συμπέρανε εμπειρικά ότι η δυναμικότητα κάθε ομάδας επηρεάζεται σημαντικά από τους τραυματισμούς, τις ανταλλαγές (trades) και άλλους παράγοντες που επιδρούν στη σύνθεση και «χημεία» της ομάδας.

Ένα σημαντικό ζήτημα της πλειοψηφίας των μελετών που έχουν πραγματοποιηθεί και προτείνουν μοντέλα πρόβλεψης αφορά την εκ των υστέρων πληροφόρηση. Σε έναν τόσο δυναμικό τομέα, που συνδέεται άμεσα με την οικονομία της χώρας στην οποία λαμβάνει χώρα μια διοργάνωση καλαθοσφαίρισης, είναι δύσκολο να αντληθούν πληροφορίες εκ των προτέρων, τόσο οι οικονομικές όσο και οι αγωνιστικές, όπως είναι εμφανές από τη φύση του αντικειμένου μελέτης. Επίσης, ο κύριος όγκος των εργασιών που έχουν πραγματοποιηθεί επικεντρώνεται στην έρευνα της συσχέτισης των νικών με μερικούς αγωνιστικούς παράγοντες. Επιπρόσθετα, στην πλειονότητα των ερευνών έχει αγνοηθεί το πρόβλημα της ενδογένειας των μεταβλητών.

2.2 Χωρικά Οικονομετρικά Υποδείγματα

Τα τελευταία χρόνια έχει σημειωθεί μεγάλη πρόοδος στο αντικείμενο της χωρικής οικονομετρίας, πάνω στο οποίο θεμελιώνεται μεθοδολογικά η παρούσα μελέτη. Ξεκινώντας από τα διαστρωματικά δεδομένα (cross-sectional), ο Lee (2004) απέδειξε τη συνέπεια και την ασυμπτωτική κανονικότητα του εκτιμητή Μεγίστης Πιθανοφάνειας (Maximum Likelihood - ML) ή Quasi-Maximum Likelihood (QML), ενώ οι Kelejian and Prucha (1998, 1999) ανέπτυξαν έναν εκτιμητή της Γενικευμένης Μεθόδου Ροπών (GMM) και τις ασυμπτωτικές του ιδιότητες για ένα μοντέλο

παλινδρόμησης που περιλαμβάνει χωρικές αυτο-παλίνδρομες διαταραχές. Εκτός αυτού, ο Lee (2007) κατέληξε στον καλύτερο εκτιμητή GMM και έδειξε ότι προσεγγίζεται από την ίδια κατανομή με τον εκτιμητή ML (υπό κανονικές υποθέσεις). Οι Kelejian and Prucha (2010) και οι Lin and Lee (2010) ανέπτυξαν επίσης εκτιμητές GMM λαμβάνοντας υπόψη την ύπαρξη ετεροσκεδαστικότητας.

Ωστόσο, δεν περνά απαρατήρητο το ενδιαφέρον της χωρικής οικονομετρικής βιβλιογραφίας για τον προσδιορισμό και την εκτίμηση οικονομετρικών σχέσεων με βάση χωρικά πάνελ δεδομένα. Αυτό το ενδιαφέρον μπορεί να εξηγηθεί, εν μέρει, από την αυξημένη διαθεσιμότητα περισσότερων συνόλων δεδομένων καθώς και από το γεγονός ότι τα δεδομένα της μορφής αυτής προσφέρουν εκτεταμένες δυνατότητες μοντελοποίησης σε σύγκριση με τα διαστρωματικά δεδομένα. Επιπλέον, τα δεδομένα panel παρέχουν γενικά περισσότερη πληροφόρηση, περιέχουν περισσότερες παραλλαγές και μικρότερη συγγραμμικότητα μεταξύ των μεταβλητών. Η χρήση δεδομένων panel έχει ως αποτέλεσμα περισσότερους βαθμούς ελευθερίας που συνεπάγονται με τη σειρά τους αποτελεσματικότερη εκτίμηση.

Αρκετοί συγγραφείς εστίασαν την έρευνά τους στην ανάπτυξη εκτιμητών στατικών και δυναμικών μοντέλων χωρικής διάχυσης για δεδομένα panel. Έχουν αναπτυχθεί τρεις μέθοδοι στη βιβλιογραφία για την εκτίμηση μοντέλων διάχυσης ως προς το χώρο και το χρόνο. Μια μέθοδος είναι ο εκτιμητής Μεγίστης Πιθανοφάνειας (ML) διορθωμένος όσον αφορά τη μεροληψία ή Quasi-Μέγιστης Πιθανότητας (QML). Ακολουθούν η Γενικευμένη Μέθοδος Ροπών (IV/GMM) που βασίζεται σε βοηθητικές μεταβλητές (instrumental variables) και η Bayesian Markov Chain Monte Carlo (MCMC). Έχουν πραγματοποιηθεί ποικίλες έρευνες και διατριβές βασιζόμενες και στις τρεις αυτές μεθόδους.

Αρχικά, η επέκταση του χωρικού οικονομετρικού μοντέλου σε ένα χωροχρονικό υπόδειγμα για μια ομάδα παρατηρήσεων N με την πάροδο του χρόνου T περιόδων πραγματοποιείται με την προσθήκη ενός δείκτη t στο απλούστερο μοντέλο, ο

οποίος τρέχει από 1 έως T, στις μεταβλητές και τον όρο σφάλματος αυτού του μοντέλου:

$$Y_t = \rho W Y_t + \alpha I_N + X_t\beta + W X_t \theta + u_t \text{ , όπου } u_t = \lambda W u_t + \varepsilon_t \qquad (2.1)$$

Σημαντικό μειονέκτημα αποτελεί το γεγονός ότι το εκτιμώμενο μοντέλο δεν λαμβάνει υπόψη τη χωρική και χρονική ετερογένεια. Οι χωρικές μονάδες είναι πιθανό να διαφέρουν ως προς τις μεταβλητές που επηρεάζουν την εξαρτημένη μεταβλητή αλλά που είναι δύσκολο να μετρηθούν. Η αποτυχία μέτρησης αυτών των μεταβλητών αυξάνει τον κίνδυνο λήψης μεροληπτικών αποτελεσμάτων εκτίμησης. Μια λύση είναι να εισαχθεί στο υπόδειγμα μια μεταβλητή σταθερή στο χρόνο (intercept) που αντιπροσωπεύει την επίδραση των παραλειπόμενων μεταβλητών που είναι ιδιόρρυθμες σε κάθε χωρική μονάδα.

Η εκτίμηση των στατικών μοντέλων χωρικής διάχυσης panel δεδομένων συζητείται εκτενώς από τους Elhorst (2003, 2010a), Anselin et al. (2007) και Lee and Yu (2010a). Ο Elhorst (2003, 2010a) παρουσιάζει τον εκτιμητή ML του υποδείγματος χωρικής υστέρησης και του μοντέλου σφάλματος έχοντας συμπεριλάβει σταθερές ή τυχαίες επιδράσεις. Για το μοντέλο σταθερών επιδράσεων, οι παράμετροι απόκρισής του μπορούν να εκτιμηθούν συγκεντρώνοντας πρώτα τις σταθερές επιδράσεις, κατά τους Baltagi (2005) και Elhorst (2010a). Η παραγόμενη εξίσωση μπορεί στη συνέχεια να εκτιμηθεί από τη διαδικασία ML που αναπτύχθηκε από την Anselin (1988) για το μοντέλο χωρικής υστέρησης, υπό την προϋπόθεση ότι αυτή η διαδικασία γενικεύεται από μία μόνο διατομή N παρατηρήσεων σε T διατομές από N παρατηρήσεις. Οι Lee and Yu (2010a) έδειξαν ότι ο εκτιμητής ML της χωρικής υστέρησης και του μοντέλου χωρικής διάχυσης του σφάλματος με χωρικές σταθερές επιδράσεις, όπως ορίζεται από τον Elhorst (2003, 2010a), θα αποδώσει μια ασυνεπή εκτίμηση παραμέτρων της διακύμανσης εάν το N είναι μεγάλο και το T μικρό, αλλά και ασυνεπείς εκτιμήσεις όλων των παραμέτρων της χωρικής υστέρησης και του μοντέλου χωρικού σφάλματος με χωρικές και χρονικές σταθερές επιδράσεις εάν και τα δύο, N και T, είναι μεγάλα.

Με την πάροδο του χρόνου παρατηρείται αυξανόμενη βιβλιογραφία αφιερωμένη σε δυναμικά χωρικά panel. Χαρακτηριστικά είναι τα έργα των Elhorst (2012, 2014)

και Lee and Yu (2010, 2015). Οι δυναμικοί χωρικοί πίνακες είναι σε θέση να αντιμετωπίσουν μη παρατηρήσιμες χωρικές, ατομικές ή και ειδικές χρονικές επιδράσεις. Αντιμετωπίζουν επίσης πιο αποτελεσματικά τα προβλήματα ενδογένειας, όπως η πιθανή μεροληψία στον συντελεστή της χωρικής υστέρησης της εξαρτημένης μεταβλητής.

Το μοντέλο χωρικής διάχυσης με δεδομένα panel ανάγεται σε δυναμικό αν προστεθούν χρονικές υστερήσεις των μεταβλητών Y_t και WY_t :

$$Y_t = \tau Y_{t-1} + \rho W Y_t + \eta W Y_{t-1} + \alpha I_N + X_t \beta + W X_t \theta + \mu + \xi_t I_N + u_t \qquad (2.2)$$

Αυτό το υπόδειγμα είναι γνωστό ως δυναμικό μοντέλο χωρικής διάχυσης Durbin (Debarsy et al., 2011). Ομοίως, μπορούν να προστεθούν χρονικές υστερήσεις των μεταβλητών X_t και WX_t και των όρων σφάλματος u_t και Wu_t. Ο Elhorst (2012) δίνει μια επισκόπηση των κύριων περιορισμών που έχουν εξεταστεί στη βιβλιογραφία: (i) $\theta=0$ για να αποκλειστούν οι εξωγενείς επιδράσεις αλληλεπίδρασης (WX_t), (ii) $\rho=0$ για να αποκλειστούν οι ταυτόχρονες επιδράσεις ενδογενούς αλληλεπίδρασης (WY_t), (iii) $\eta=0$ για να αποκλειστούν οι υστερήσεις της ενδογενούς αλληλεπίδρασης των επιδράσεις (WY_{t-1}), και (iv) $\eta=-\tau\rho$.

Οι Baltagi et al. (2014) προτείνουν ένα δυναμικό χωρικό μοντέλο που επιτρέπει τη χρονική καθώς και χωρική εξάρτηση, αλλά περιορίζει τη συνδιακύμανση χρόνου-χώρου στο μηδέν ακολουθώντας ορισμένες προγενέστερες μελέτες, όπως αυτές των Franzese and Hays (2007), Kukenova and Monteiro (2009), Jacobs et al. (2009), Korniotis (2010), Elhorst (2010) και Brady (2011). Επικεντρώνονται σε διάφορους εκτιμητές για το δυναμικό και αυτοπαλίνδρομο χωρικής υστέρησης υπόδειγμα panel δεδομένων με χωρικά συσχετισμένες διαταραχές.

Στο πνεύμα των Arellano and Bond (1991) και Mutl (2006), προτείνεται μια δυναμική χωρική γενικευμένη μέθοδος ροπών (GMM), συμπεριλαμβανομένης της προσέγγιση των Kapoor et al. (2007) για τον χωρικά αυτοπαλίνδρομο όρο σφάλματος (Spatial AutoRegressive-SAR). Ο υπολειπόμενος όρος του τελευταίου θεωρείται ότι έχει δομή τυχαίων επιδράσεων. Να αναφερθεί στο σημείο αυτό ότι οι Kapoor et al. (2007) ήταν οι πρώτοι που εξήγαν έναν εκτιμητή GMM για ένα μοντέλο με “one-way” χωρικά αυτο-παλίνδρομο σφάλμα σε panel δεδομένα. Οι

Mutl and Pfaffermayr (2011) επέκτειναν την παραπάνω μελέτη, λαμβάνοντας υπόψη την ύπαρξη ενδογενούς χωρικής υστέρησης και παρήγαγαν ένα στατιστικό στοιχείο Hausman για να ελέγξει την «επάρκεια» της προδιαγραφής των τυχαίων επιδράσεων σε αυτήν την προσέγγιση GMM.

Μια γενικότερη προδιαγραφή του dynamic spatial panel data model έχει προταθεί από τους Yu et al. (2008, 2012) και Lee and Yu (2010, 2014) που εφαρμόζουν ένα μοντέλο που επιτρέπει την χρονική και χωρική εξάρτηση καθώς και την ανάμειξη συνιστωσών με χωροχρονική εξάρτηση που μπορεί να ερμηνευτεί ως χωρική διάχυση που εξαπλώνεται με την πάροδο του χρόνου.

Πιο συγκεκριμένα, οι Yu et al. (2008) κατασκεύασαν έναν εκτιμητή διορθωμένης μεροληψίας για ένα δυναμικό μοντέλο (Y_{t-1}, WY_t και WY_{t-1}) με χωρικές σταθερές επιδράσεις. Οι Lee and Yu (2010b) επεκτείνουν αυτή τη μελέτη για να συμπεριλάβουν σταθερές επιδράσεις της χρονοσειράς. Πρώτα, εκτιμούν το μοντέλο από τον εκτιμητή ML που αναπτύχθηκε από τον Elhorst (2003, 2010a) για το μοντέλο χωρικής υστέρησης με χωρικές (και χρονικές) σταθερές επιδράσεις, ο οποίος βασίζεται στην υπό συνθήκες λογαριθμοποιημένη συνάρτηση πιθανοφάνειας του μοντέλου. Στη συνέχεια, παρέχουν μια αυστηρή ασυμπτωτική θεωρία για τον εκτιμητή και η μεροληψία τους διορθώθηκε όταν και ο αριθμός των χωρικών μονάδων (N) και του αριθμού των χρονικών περιόδων (T) στο δείγμα τείνουν στο άπειρο έτσι ώστε το όριο μεταξύ N και T υπάρχει και οριοθετείται μεταξύ μηδέν και άπειρου ($0 < lim(N/T) < \infty$). Σύμφωνα με τους Lee and Yu (2010c), αυτή η προϋπόθεση υποδηλώνει ότι T -> ∞ , όπου T δεν μπορεί να είναι πολύ μικρό σε σχέση με το N. Η διόρθωση μεροληψίας προκύπτει τόσο για κανονικά κατανεμημένους όρους σφάλματος (ML) όσο και για όρους σφάλματος που δεν βασίζονται στην υπόθεση κανονικότητας (QML). Ο Elhorst (2010c) διερευνά τις μικρές ιδιότητες δείγματος αυτού του διορθωμένου ως προς τη μεροληψία εκτιμητή, επεκτείνοντας τον άνευ όρων εκτιμητή ML που προτείνεται από τους Hsiao et al. (2002) με το μεταβλητό πλάτος. Για τον προσδιορισμό της αναμενόμενης τιμής και της διακύμανσης εφαρμόζει την προσέγγιση Bhargava και Sargan (1983). Ένα από τα συμπεράσματά του είναι ότι η εκτίμηση παραμέτρων της μεταβλητής WY_t εξακολουθεί να είναι σημαντικά μεροληπτική όταν

χρησιμοποιείται αυτός ο άνευ όρων εκτιμητής ML. Ωστόσο, ανάγει έναν μικτό εκτιμητή καλύτερο από αυτόν των Yu et al. (2008) για μικρές τιμές *T* (*T*=5).

Οι Lee and Yu (2010c, e) ενσωμάτωσαν τις πρόσφατες εξελίξεις στα δυναμικά υποδείγματα χωροχρόνου και εκτίμησαν ένα στατιστικό ελέγχου Hausman και πολλαπλασιαστή Lagrange για τον έλεγχο της συνάφειας των τυχαίων επιδράσεων στο στατικό μοντέλο χωρικής διάχυσης στο πλαίσιο της μεγίστης πιθανοφάνειας.

Οι Parent and LeSage (2010, 2011, 2012) και Debarsy et al. (2012) επεκτείνουν την προσέγγιση των Yu et al. (2008,2012), Lee και Yu (2010, 2014) εισάγοντας ένα χωροχρονικό «φίλτρο» που υποδηλώνει έναν περιορισμό στον όρο ανάμειξης που αντικατοπτρίζει τη χωρική διάχυση στην πάροδο του χρόνου. Αυτός ο τύπος περιορισμού προκαλεί μια διαχωριστικότητα της εξάρτησης χρόνου και χώρου που απλοποιεί διαδικασία εκτίμησης, ειδικά στην προσέγγιση με Bayesian Markov Chain Monte Carlo (MCMC).

Οι Lee και ο Yu (2016) θεωρούν ότι ένα χωρικό Durbin δυναμικό μοντέλο που περιλαμβάνει ταυτόχρονα χρονική, χωρική και χωροχρονική επίδραση από τις επεξηγηματικές μεταβλητές και δείχνουν ότι οι παράμετροι γενικά προσδιορίζονται μέσω της Μεθόδου των Ελαχίστων Τετραγώνων σε δύο στάδια (2 Stage Least Squares - 2sls).

Συνεχίζοντας την επισκόπηση, οι Baltagi et al. (2018) προτείνουν έναν χωρικό εκτιμητή GMM βασιζόμενοι στην έρευνα των Arellano and Bond (1991), Mutl (2006) και Kapoor et al. (2007). Χρησιμοποιώντας τεχνικές Monte Carlo, συγκρίνουν την εμπειρική απόδοση ενός GM χωρικού εκτιμητή με αυτόν της OLS, τον εκτιμητή GMM των Arellano and Bond (1991), πρόκειται για εκτιμητές δεδομένων panel που δεν λαμβάνουν υπόψη τη χωρική δομή των διαταραχών, ειδικά σε βραχυπρόθεσμες και μακροπρόθεσμες επιδράσεις, και με μη καθορισμένους χωρικούς εκτιμητές GM όπως αυτόν του Mutl (2006). Οι Baltagi et al. (2018) συμπεριλαμβάνουν στο υπόδειγμα μια χρονική υστέρηση (καταγραφή της χρονικής εξάρτησης), μια χωρική υστέρηση (που αντιπροσωπεύει τη χωρική εξάρτηση), και έναν όρο που εκφράζει τη χωροχρονικά αποτελέσματα διάχυσης. Πολλές θεωρητικές ή / και εφαρμοσμένες διατριβές βασίζονται σε αυτή τη δομή, με βάση μια προγενέστερη θεωρία ως

πρότυπο που δεν «απαιτεί» για WXs στο μοντέλο. Ενδεικτικά, αυτό αναφέρεται στους Yu et al. (2008), Parent and LeSage (2010, 2011, 2012), Lee και Yu (2014) ή Yang (2017) μεταξύ άλλων.

Παρά τη βελτίωση των γνώσεων στα θεωρητικά θεμέλια της χωρικής οικονομετρίας, εξακολουθεί να υπάρχει σύγχυση σχετικά με την ερμηνεία των επιπτώσεων μιας αλλαγής σε μια επεξηγηματική μεταβλητή για την εξαρτώμενη μεταβλητή στο χωρικό αυτό-παλίνδρομο μοντέλο. Με την ανάπτυξη του μοντέλου χωρικής διάχυσης δεδομένων panel, και πιο συγκεκριμένα, χωροχρονικής διάχυσης, όπου αλληλεπιδρούν ο χρόνος και οι χωρικές διαστάσεις, η ανάγκη για σαφείς ερμηνείες των επιπτώσεων της αλλαγής στις επεξηγηματικές μεταβλητές είναι ισχυρότερη από ποτέ. Τα δυναμικά χωρικά μοντέλα είναι χρήσιμα σε πολλούς τομείς της οικονομικής ανάλυσης αποτελώντας ένα πλήρες εργαλείο για την ανάλυση των επιδράσεων διάχυσης (χωρικές, χρονικές και χωροχρονικές).

Οι Fingleton and Szumilo (2019) ανέπτυξαν ένα δυναμικό υπόδειγμα χωρικής διάχυσης δεδομένων panel, όπως προτείνεται από τους Baltagi et al. (2014, 2018). Σε μια προσπάθεια αποφυγής των ad hoc υποθέσεων το μοντέλο περιλαμβάνει χρονικές και χωρικές υστερήσεις. Αυτό διαφοροποιεί την προσέγγισή τους από απλούστερες μεθόδους που χρησιμοποιούνται στη βιβλιογραφία. Η επιλογή μιας διαδικασίας σφάλματος spatial moving average επιχειρεί να μετριάσει πιθανές επιβλαβείς επιδράσεις για την εκτίμηση των εξωγενών μεταβλητών με υστέρηση, ενώ με τη στρατηγική των instruments (βοηθητικές μεταβλητές) επιδιώκεται η εξάλειψη της μεροληψίας λόγω πιθανής ενδογένειας.

2.3 Συνεισφορά στη βιβλιογραφία

Στην παρούσα εργασία εκτιμάται ένα δυναμικό μοντέλο χωρικής διάχυσης Durbin σε δεδομένα panel για στοιχεία που αφορούν ένα άθλημα ομαδικό και πηγάζουν από τη μεγαλύτερη, καλύτερα οργανωμένη και οικονομικά αναπτυγμένη μπασκετική διοργάνωση του NBA. Πρόκειται, λοιπόν, για ένα αντικείμενο μελέτης με ιδιαίτερο ενδιαφέρον, εφόσον στο πλαίσιο συχνών μεταβολών οι διοικήσεις των ομάδων καλούνται να λαμβάνουν αποφάσεις στοχεύοντας στην επίτευξη του καλύτερου δυνατού αποτελέσματος σε ένα ιδιαίτερα ανταγωνιστικό και δυναμικό περιβάλλον.

Στο πλαίσιο των ζητημάτων για τα οποία γίνεται λόγος στο τέλος της πρώτης παραγράφου του κεφαλαίου αυτού, κατ' αρχάς, στη συγκεκριμένη μελέτη πέρα από τους αγωνιστικούς παράγοντες εξετάζονται και παράγοντες που σχετίζονται με τις διοικητικές και οικονομικές αποφάσεις κάθε ομάδας και, συγκεκριμένα, τις ανταλλαγές, την επιλογή για το draft και την οικονομική πολιτική ανάλογα με το τι μέρος του salary cap επιλέγει να δεσμεύσει. Αυτές οι μη αγωνιστικές μεταβλητές είναι ικανές να συμβάλουν σε έναν βαθμό και στην πρόβλεψη, καθώς ανάλογα με το επίπεδο μιας ομάδας και τους στόχους που θέτει λαμβάνει και τις κατάλληλες αποφάσεις.

Ως προς τη μεθοδολογία, η συνεισφορά της παρούσας έρευνας στη βιβλιογραφία, συγκριτικά με προηγούμενες μελέτες πάνω σε σχετικά αντικείμενα, αφορά στο γεγονός ότι το μοντέλο που εκτιμάται περιλαμβάνει χωρική διάχυση των μη αγωνιστικών επεξηγηματικών μεταβλητών, χρονική υστέρηση καθώς και χωροχρονική υστέρηση της εξαρτημένης μεταβλητής που εκφράζει τις νίκες μιας ομάδας. Με την ανάπτυξη ενός τέτοιου υποδείγματος είναι δυνατό να εξεταστεί κατά πόσο επηρεάζονται τα αγωνιστικά αποτελέσματα μιας ομάδας από τις αποφάσεις και το ρεκόρ νικών-ηττών, κατά την προηγούμενη αγωνιστική περίοδο, τόσο της ίδιας της ομάδας όσο και των αντιπάλων της. Επίσης, επιλέχθηκε η Μέθοδος Μεγίστης Πιθανοφάνειας για την εκτίμηση του δυναμικού υποδείγματος, προκειμένου να αποφευχθούν τα προβλήματα ενδογένειας των ανεξάρτητων

μεταβλητών, κάτι που δεν έχει ληφθεί υπόψη κατά κύριο λόγο στην έως τώρα βιβλιογραφία.

Η χωρική οικονομετρία δεν συνηθίζεται στον τομέα της αθλητικής στατιστικής. Ωστόσο αποτελεί «ικανό» εργαλείο, αν αναλογιστούμε ότι χρησιμοποιείται, πλέον, ευρέως σε πολλούς κλάδους των οικονομικών.

Κεφάλαιο 3: Μεθοδολογικό πλαίσιο

3.1 Γενικευμένο υπόδειγμα

3.1.1 Το Υπόδειγμα Χωρικής Διάχυσης Durbin (Spatial Durbin Model)

Η μεθοδολογική προσέγγιση έχει ως αφετηρία το γραμμικό μοντέλο παλινδρόμησης πάνελ δεδομένων, της μορφής:

$$y_{it} = \alpha + x_{it}\,\beta + \mu_i + \varepsilon_{it} \tag{3.1}$$

όπου ο δείκτης *i* δηλώνει τη διαστρωματική μονάδα[18] (*i = 1, ..., N*) και ο δείκτης *t* τον χρόνο *(t = 1, ..., T)* , x_{it} είναι ένα *1×k* διάνυσμα παρατηρήσεων των ανεξάρτητων μεταβλητών, β είναι το *k×1* διάνυσμα των σταθερών προς εκτίμηση παραμέτρων, ε_{it} είναι ο ανεξάρτητος και πανομοιότυπα κατανεμημένος όρος του σφάλματος, με μηδενική μέση τιμή και σταθερή διακύμανση, ενώ η μεταβλητή μ_i αποτυπώνει τη διαστρωματική διαφοροποίηση, και ελέγχει όλους τους χρονικά αμετάβλητους παράγοντες που επηρεάζουν την εξαρτημένη μεταβλητή και διαφέρουν μεταξύ των υπό μελέτη ομάδων. Η παράλειψή τους από το υπόδειγμα θα οδηγούσε σε μεροληπτικές εκτιμήσεις σε μια τυπική μελέτη με δεδομένα panel (Baltagi, 2005).

Σε μορφή πινάκων το υπόδειγμα γράφεται ισοδύναμα:

$$Y_t = a + X_t\,\beta + \mu + \varepsilon_{it} \tag{3.2}$$

Υπάρχουν τρεις διαφορετικοί τύποι επιδράσεων αλληλεπίδρασης που δύνανται να εξηγήσουν γιατί μια παρατήρηση που σχετίζεται με μια συγκεκριμένη ομάδα (διαστρωματική μονάδα) μπορεί να εξαρτάται από παρατηρήσεις σε άλλες ομάδες:

α) ενδογενείς αλληλεπιδράσεις που αφορούν στην εξαρτημένη μεταβλητή,

β) εξωγενείς αλληλεπιδράσεις που αφορούν στις ανεξάρτητες μεταβλητές και

[18] Στην παρούσα εφαρμογή, η διαστρωματική μονάδα είναι η κάθε ομάδα του πρωταθλήματος του NBA: i = 1,…,30.

γ) αλληλεπιδράσεις που αφορούν στον όρο του σφάλματος.

Λαμβάνοντας υπόψη και τα παραπάνω τρία είδη αλληλεπιδράσεων, το υπόδειγμα της εξίσωσης (3.1) μπορεί να επεκταθεί ως εξής:

$$y_{it} = \alpha + \rho \sum_{j=1}^{N} w_{ij}\, y_{jt} + x_{it}\, \beta + \theta \sum_{j=1}^{N} w_{ij}\, x_{ijt} + \mu_i + \varepsilon_{it} \qquad (3.3)$$

όπου: $\varepsilon_{it} = \lambda \sum_{j=1}^{N} w_{ij}\, \varepsilon_{jt} + u_{it}$.

Ισοδύναμα με μορφή πινάκων: $Y = \rho WY + bX + qWX + \varepsilon + aI_N$ (3.4)

όπου $\varepsilon = l_N W\varepsilon + u$.

Το μοντέλο αυτό περιλαμβάνει:

(α) μία μεταβλητή χωρικής υστέρησης (spatial lag) της εξαρτημένης μεταβλητής $\sum_{j=1}^{N} w_{ij}\, y_{jt}$,

(β) τις χωρικές υστερήσεις των ανεξάρτητων μεταβλητών $\sum_{j=1}^{N} w_{ij}\, x_{ijt}$,και

(γ) μια χωρικά αυτοπαλίνδρομη διαδικασία για τον όρο του σφάλματος $\lambda \sum_{j=1}^{N} w_{ij}\, \varepsilon_{jt}$ (spatial error).

Η παράμετρος ρ της χωρικής υστέρησης της εξαρτημένης μεταβλητής καλείται συντελεστής χωρικής αυτοπαλινδρόμησης (spatial autoregressive coefficient), η παράμετρος λ συντελεστής χωρικής αυτοσυσχέτισης (spatial autocorrelation coefficient), ενώ ο συντελεστής θ εκφράζει την ύπαρξη επιδράσεων διάχυσης των επεξηγηματικών μεταβλητών.

Τέλος, το w_{ij} αποτελεί ένα στοιχείο μίας θετικής, τετραγωνικής, χωρικής μήτρας W βαρών που περιγράφει την σχέση[19] των διαστρωματικών μονάδων του δείγματος.

Παρά το γεγονός ότι ίσως είναι αποτελεσματικότερο να συμπεριληφθούν στο υπόδειγμα (3.3) και οι τρεις επιδράσεις αλληλεπίδρασης ταυτόχρονα, ο Manski (1993) υποστηρίζει την ανάγκη να αγνοηθεί τουλάχιστον μια εκ των τριών σχέσεων αλληλεπίδρασης, προκειμένου να καταστεί εφικτός ο προσδιορισμός των παραμέτρων.

[19] Στη συγκεκριμένη εφαρμογή, η σχέση των διαστρωματικών μονάδων οφείλεται κυρίως στην οργανωσιακή δομή του πεδίου μελέτης (NBA).

Για τον σκοπό αυτό, διακρίνονται τέσσερεις υποπεριπτώσεις του αρχικού υποδείγματος: α) θέτοντας $\lambda=0$ και καλείται χωρικό υπόδειγμα του Durbin (spatial Durbin model), β) θέτοντας $\theta=0$ και καλείται υπόδειγμα χωρικής αυτοσυσχέτισης (spatial autocorrelation model), γ) θέτοντας $\lambda=0$ και $\theta=0$ και προκύπτει ένα χωρικά αυτοπαλίνδρομο υπόδειγμα (spatial autoregressive model) και, δ) θέτοντας $\rho=0$ και $\theta=0$ το οποίο οδηγεί σε ένα υπόδειγμα χωρικού σφάλματος (spatial error model).

Το χωρικό μοντέλο Durbin θεωρείται το καταλληλότερο για την εισαγωγή των χωρικών επιδράσεων, καθώς λαμβάνει υπόψη του τις επιδράσεις χωρικής διάχυσης των υπό εξέταση ανεξάρτητων μεταβλητών (LeSage και Pace, 2009, Elhorst, 2010) και παριστάνεται ως εξής:

$$y_{it} = \rho \sum_{j=1}^{N} w_{ij}\, y_{jt} + x_{it}\, \beta + \sum_{j=1}^{N} w_{ij}\, x_{ijt}\, \gamma + \mu_i + u_{it} \qquad (3.5)$$

όπου β και γ είναι διανύσματα $k\times1$ σταθερών αλλά άγνωστων παραμέτρων.

3.1.2 Το Δυναμικό Υπόδειγμα Χωρικής Διάχυσης Durbin (Dynamic Spatial Durbin Model ή D-SDM)

Σύμφωνα με τους Fingleton και Szumilo (2019), μπορούμε να συμπεριλάβουμε επιπλέον στο υπόδειγμα (3.5) και χρονική υστέρηση (temporal lag) της εξαρτημένης μεταβλητής, ώστε να αναχθεί σε δυναμικό (όσον αφορά την εξαρτημένη μεταβλητή). Προκειμένου να ληφθεί υπόψη η χωρική αλληλεπίδραση μεταξύ διαστρωματικών μονάδων (δηλαδή των ομάδων), στο μοντέλο (3.5) προστίθενται: (α) επιδράσεις ενδογενούς αλληλεπίδρασης με την εξαρτημένη μεταβλητή, (β) επιδράσεις εξωγενούς αλληλεπίδρασης με τη χρονική υστέρηση της εξαρτημένης μεταβλητής και (γ) επιδράσεις εξωγενούς αλληλεπίδρασης με τις ανεξάρτητες μεταβλητές.

Συνεπώς, το Δυναμικό Υπόδειγμα Χωρικής Διάχυσης Durbin (Dynamic Spatial Durbin Model ή D-SDM) έχει τη μορφή:

$$y_{it} = at + \tau y_{it-1} + \rho \sum_{j=1}^{N} w_{ij}\, y_{jt} + x_{it}\, \beta + \eta \sum_{j=1}^{N} w_{ij}\, y_{jt-1} + \theta \sum_{j=1}^{N} w_{ij}\, x_{ijt} + \mu_i + u_{it} \qquad (3.6)$$

όπου: y_{it-1} η χρονική υστέρηση (temporal lag), $\sum_{j=1}^{N} w_{ij}\, y_{jt}$ η χωρική διάχυση (spatial lag) και $\sum_{j=1}^{N} w_{ij}\, y_{jt-1}$ η χωρο-χρονική υστέρηση (spatiotemporal lag) της εξαρτημένης μεταβλητής ενώ $\sum_{j=1}^{N} w_{ij}\, x_{ijt}$ είναι η χωρική διάχυση (spatial lag) των ανεξάρτητων μεταβλητών.

Η ισοδύναμη μορφή πινάκων του υποδείγματος είναι η εξής:

$$Y_t = at + \tau Y_{t-1} + \rho W Y_t + \beta X_t + \eta W Y_{t-1} + \theta\, W X_t + M + \Delta F + u_t$$

ή (3.7)

$$Y_t = (I - \rho W)^{-1}(\tau I + \eta W) Y_{t-1} + (I - \rho W)^{-1}(\beta + \theta\, W) X_t + (I - \rho W)^{-1}(at + M + \Delta F + u_t)$$

3.1.3 Άμεσες, Έμμεσες και Συνολικές Επιδράσεις

Δεδομένου ότι το D-SDM εκμεταλλεύεται την αλληλεξαρτώμενη σχέση μεταξύ των ομάδων, η επίδραση μιας εξωγενούς αλλαγής σε μια επεξηγηματική μεταβλητή μιας συγκεκριμένης ομάδας, επηρεάζει όχι μόνο την ίδια την ομάδα, αλλά και όλες τις υπόλοιπες ομάδες, έμμεσα. Προκειμένου να εκτιμηθούν τα χωρικά αποτελέσματα διάχυσης με βάση την εκτίμηση του D-SDM, υπολογίζονται οι άμεσες, έμμεσες και συνολικές επιδράσεις (LeSage and Pace, 2009). Τα μέτρα αυτά αποτυπώνουν το σωρευτικό/αθροιστικό αποτέλεσμα των αλλαγών στις ανεξάρτητες μεταβλητές, οι οποίες οδηγούν σε μεταβολή της μακροπρόθεσμης ισορροπίας σταθερής κατάστασης (Yu et al., 2013).

Σε ένα δυναμικό SDM, η παράγωγος του *y* ως προς το x_k δεν είναι ίση με τον συντελεστή β_k αλλά με:

$$\Gamma = \frac{\partial y}{\partial x_k} = (I - \rho W)^{-1} \begin{bmatrix} \beta_k & w_{12}\theta_k & \dots & w_{1n}\theta_k \\ w_{21}\theta_k & \beta_k & \dots & w_{2n}\theta_k \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}\theta_k & w_{n2}\theta_k & \dots & \beta_k \end{bmatrix} \tag{3.8}$$

όπου για κάθε ομάδα και επεξηγηματική μεταβλητή τα διαγώνια στοιχεία του Γ ($\gamma_{ii} \equiv \beta_k (I - \rho W)^{-1}$) αντιστοιχούν στην άμεση επίδραση ενώ τα μη διαγώνια στοιχεία ($\gamma_{ij} \equiv w_{ij}\, \beta_k (I - \rho W)^{-1}$, with $i \neq j$) στην έμμεση επίδραση. Σύμφωνα με τους LeSage and Pace

(2009) οι άμεσες, έμμεσες και συνολικές επιδράσεις μπορούν να υπολογιστούν όπως φαίνεται ακολούθως:

$$\bar{M}(k)_{total} = n^{-1} l'_n \Gamma l_n$$

$$\bar{M}(k)_{direct} = n^{-1} trace(\Gamma) \qquad (3.9)$$

$$\bar{M}(k)_{indirect} = \bar{M}(k)_{total} - \bar{M}(k)_{direct}$$

όπου l_n είναι ένα $n \times 1$ μοναδιαίο διάνυσμα. Με αυτό τον τρόπο, είναι δυνατό να υπολογιστούν τόσο το σημείο όσο και το μέγεθος των άμεσων και έμμεσων επιδράσεων των επεξηγηματικών μεταβλητών.

3.1.4 Η Μέθοδος της Μέγιστης Πιθανοφάνειας σε Δεδομένα Panel[20]

Προκειμένου να αντιμετωπιστεί το πρόβλημα της ενδογένειας του υποδείγματος, εφαρμόζεται η μέθοδος Μέγιστης Πιθανοφάνειας (Maximum Likelihood - ML). Στη μέθοδο ML οι χωρικές επιδράσεις *(μ_{it})* μπορούν να αντιμετωπιστούν ως σταθερές (fixed specific effects) ή τυχαίες (random specific effects). Στο μοντέλο σταθερών επιδράσεων εισάγεται μία ψευδομεταβλητή (dummy variable) για κάθε χωρική μονάδα, ενώ στο μοντέλο τυχαίων επιδράσεων η μ_{it} αντιμετωπίζεται ως μία τυχαία μεταβλητή με μηδενική μέση τιμή και σταθερή διακύμανση. Επιπλέον, θεωρείται ότι οι τυχαίες μεταβλητές μ_{it} και ε_{it} είναι ανεξάρτητες μεταξύ τους.

Ας μελετήσουμε την εκτίμηση του χωρικού μοντέλου Durbin (χωρίς χρονικές υστερήσεις) με τη μέθοδο Μέγιστης Πιθανοφάνειας. Το προς μελέτη υπόδειγμα είναι το εξής:

$$y_{i,t} = a + \rho \sum_{j=1}^{N} w_{ij}\, y_{j,t} + \sum_{c=1}^{k} \beta_c x_{c,i,t} + \sum_{c=1}^{k} \theta_c \sum_{j=1}^{N} w_{ij} x_{c,j,t} + \mu_i + e_{i,t} \qquad (3.10)$$

- Η περίπτωση των σταθερών επιδράσεων μ_i

Ένα πρώτο βήμα είναι η απαλοιφή του όρου μ_i μετασχηματίζοντας κατάλληλα τις εξαρτημένες και ανεξάρτητες μεταβλητές του υποδείγματος. Ο μετασχηματισμός των μεταβλητών πραγματοποιείται αφαιρώντας την τιμή κάθε μεταβλητής από τον «ενδο-ομαδικό» διαχρονικό της μέσο όρο, δηλαδή:

[20] Η εν λόγω ενότητα βασίζεται στον Μαρίνο (2020).

$$y^*_{i,t} = y_{i,t} - \frac{1}{T}\sum_{t=1}^{T} y_{i,t} \tag{3.11}$$

και

$$x^*_{c,i,t} = x_{c,i,t} - \frac{1}{T}\sum_{t=1}^{T} x_{c,i,t} \tag{3.12}$$

και η (3.10) γράφεται ως εξής:

$$y^*_{i,t} = a + \rho \sum_{j=1}^{N} w_{ij} y^*_{j,t} + \sum_{c=1}^{k} \beta_c\, x^*_{c,i,t} + \sum_{c=1}^{k} \theta_c \sum_{j=1}^{N} w_{ij}\, x^*_{c,j,t} + e_{i,t} \tag{3.13}$$

ή σε συνεπτυγμένη μορφή πινάκων:

$$Y_t^* = \rho\, W Y_t^* + Z_t^*\, \beta + e_t \tag{3.14}$$

και ισοδύναμα:

$$e_t = (I_N - \rho W) Y_t^* - Z_t^*\, \beta \tag{3.15}$$

όπου $Y_t^* = [y^*_{1,t}, \ldots, y^*_{N,t}]^T$, διαστάσεων $N \times 1$, $\beta = [\alpha, \beta_1, \ldots, \beta_k, \theta_1, \ldots, \theta_k]^T$ διαστάσεων

$(1+2k) \times 1$, $Z_t^* = \begin{bmatrix} 1 & x^*_{1,1,t} & \cdots & x^*_{k,1,t} & W_1 x^*_{1,1,t} & \cdots & W_1 x^*_{k,1,t} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x^*_{1,N,t} & \cdots & x^*_{k,N,t} & W_N x^*_{1,N,t} & \cdots & W_N x^*_{k,N,t} \end{bmatrix}$, διαστάσεων $N \times$

$(1+2k)$, όπου περιλαμβάνει όλες τις ανεξάρτητες μεταβλητές του υποδείγματος (σταθμισμένες και μη) ενώ β είναι το διάνυσμα των προς εκτίμηση παραμέτρων και $e_t = [e_{1,t}, \ldots, e_{N,t}]^T$, διαστάσεων $N \times 1$.

Η λογαριθμική συνάρτηση πιθανοφάνειας της εξίσωσης (3.14) είναι η ακόλουθη:

$$ln(L) = -\frac{NT}{2} ln(2\pi\sigma^2) + T\, ln|I_N - \rho W| - \frac{1}{2\sigma^2}\sum_{t=1}^{T} e_t{}' e_t \tag{3.16}$$

και με αντικατάσταση του e_t από την εξίσωση (14), γίνεται:

$$ln(L) = -\frac{NT}{2} ln(2\pi\sigma^2) + T\, ln|I_N - \rho W| - \frac{1}{2\sigma^2}\sum_{t=1}^{T}((I_N - \rho W)Y_t^* - Z_t^*\, \beta)'((I_N - \rho W)Y_t^* - Z_t^*\, \beta) \tag{3.17}$$

όπου σ^2 είναι η διακύμανση του όρου του σφάλματος και $J = |I_N - \rho W|$ είναι ο Ιακωβιανή ορίζουσα της έκφρασης του e_t ως προς Y_t^*, η οποία εισέρχεται στην εξίσωση προκειμένου να ληφθεί υπόψη η ενδογένεια του όρου $\sum_{j=1}^{N} w_{ij}\, y_{j,t}$.

Να σημειωθεί ότι η Ιακωβιανή ορίζουσα προκύπτει από την εξίσωση (3.15), παραγωγίζοντας την ως προς την εξαρτημένη μεταβλητή: $J = \left|\frac{\partial e_t}{\partial Y_t^*}\right| = |IN - \rho W|$. Η Ιακωβιανή ορίζουσα θα πρέπει να είναι θετική και πραγματική. Για να πληρούνται αυτές οι προϋποθέσεις θα πρέπει να ισχύουν οι συνθήκες ομαλότητας: $1/\omega_{min} < \rho < 1/\omega max$, οι οποίες εξασφαλίζονται με την κανονικοποίηση της W (Anselin κ.α., 2006).

Κατά τον Elhorst (2003), η διαδικασία εκτίμησης των παραμέτρων *ρ* και *β* πραγματοποιείται με βάση τον εξής επαναληπτικό αλγόριθμο:

<u>Βήμα 1</u>: Θεωρώντας τα panel δεδομένα ως διαδοχικά – *T* τον αριθμό, για *t = 1,…,T* – διαστρωματικά δεδομένα και τον *ρ* σταθερό ($\hat{\rho}$), αρχικά εκτιμώνται το διάνυσμα *β* και η διακύμανση σ^2 – συναρτήσει του $\hat{\rho}$ – μεγιστοποιώντας την εξίσωση (3.17) ως προς β και σ^2 αντίστοιχα:

$$\frac{\partial ln(L)}{\partial \beta} = 0 \Rightarrow \hat{\beta} = (Z^{*\prime}Z^*)^{-1}Z^{*\prime}(I_N - \hat{\rho}W)Y^* = (Z^{*\prime}Z^*)^{-1}Z^{*\prime}Y^* - \rho(Z^{*\prime}Z^*)^{-1}Z^{*\prime}WY^* \qquad (3.18)$$

και

$$\frac{\partial ln(L)}{\partial \sigma^2} = 0 \Rightarrow \hat{\sigma}^2 = \frac{1}{NT}((I_N - \hat{\rho}W)Y^* - Z^*\beta)'((I_N - \hat{\rho}W)Y^* - Z^*\beta)) \qquad (3.19)$$

όπου $Y^*=[Y_1^*, \ldots, Y_T^*]'$, διαστάσεων *NT×1*, και $Z^*=[Z_1^*, \ldots, Z_T^*]'$, διαστάσεων *NT×(1+ 2k)*.

Καθώς οι εκτιμητές των β και σ^2 είναι αμφότεροι συναρτήσεις του $\hat{\rho}$, η εκτίμησή τους προϋποθέτει την εκτίμηση του $\hat{\rho}$.

<u>Βήμα 2</u>: Εκτίμηση του $\hat{\rho}$: Εφαρμόζοντας την πρόταση του Anselin (1988), το πρόβλημα εκτίμησης τριών παραμέτρων μετατρέπεται σε πρόβλημα εκτίμησης μόνο μιας. Συγκεκριμένα κατά τον Anselin, με τη βελτιστοποίηση μιας συγκεντρωμένης λογαριθμικής συνάρτησης πιθανοφάνειας (concentrated log-likelihood function), συναρτήσει των εκτιμήσεων των β και σ^2 θα εκτιμηθεί το $\hat{\rho}$.

Από την εξίσωση (3.18), έπεται ότι η εκτιμήτρια του β μπορεί να αναλυθεί σε δύο επιμέρους εκτιμητές ελαχίστων τετραγώνων, μία της παλινδρόμησης της Y^* στην Z^* και μία της WY^* στη Z^*:

$$\hat{\beta} = \hat{\beta}_{OLS,Y^*\to Z^*} - \hat{\beta}_{OLS,WY^*\to Z^*} \quad (3.20)$$

όπου, κατά τα γνωστά, $\hat{\beta}_{OLS,Y^*\to Z^*} = (Z^{*\prime}Z^*)^{-1}Z^{*\prime}Y^*$ και $\hat{\beta}_{OLS,WY^*\to Z^*} = (Z^{*\prime}Z^*)^{-1}Z^{*\prime}WY^*$. Τα αντίστοιχα κατάλοιπα των δύο παλινδρομήσεων, $e_0^* = Y^* - Z^*\hat{\beta}_{OLS,Y^*\to Z^*}$ και $e_1^* = WY^* - Z^*\hat{\beta}_{OLS,WY^*\to Z^*}$, αντικαθίστανται στην (17), η οποία γίνεται:

$$\hat{\sigma}^2 = \frac{1}{NT}(e_0^* - \rho e_1^*)'(e_0^* - \rho e_1^*) \quad (3.21)$$

και αντικαθιστώντας την ακολούθως στη λογαριθμική συνάρτηση πιθανοφάνειας της εξίσωσης (3.16), προκύπτει η συγκεντρωμένη λογαριθμική συνάρτηση πιθανοφάνειας:

$$ln\,(L_c(\rho)) = C + T\,ln|I_N - \rho W| - \frac{NT}{2}(e_0 - \rho e_1)'(e_0 - \rho e_1) \quad (3.22)$$

όπου $C = -\frac{NT}{2}ln(2\pi) - \frac{NT}{2} + \frac{NT}{2}ln(NT)$, σταθερά και ανεξάρτητη της ρ. Παρατηρούμε η ότι είναι μη γραμμική συνάρτηση της ρ.

Εφόσον δεν υπάρχει κλειστή λύση στο πρόβλημα αυτό, η εκτίμηση της $\hat{\rho}$ από την (3.16) γίνεται μόνο υπολογιστικά. Εντούτοις, καθώς η $\ln(L_c)$ είναι κοίλη ως προς την δ, η λύση είναι μοναδική (Anselin and Hudak 1992, Elhorst 2003).

Για την απλοποίηση της διαδικασίας μεγιστοποίησης της $ln(L_c)$, ως προς τον όρο $ln|I_N - \rho W|$, οι Pace και Barry (1977) πρότειναν τον υπολογισμό της ορίζουσας για ένα φάσμα τιμών της δ (συνήθως ανά 0,001 μονάδες) στο διάστημα $[1/\omega_{min}, 1]$, δεδομένου ότι η W είναι κανονικοποιημένη. Με δεδομένες τις προκαθορισμένες τιμές της Ιακωβιανής ορίζουσας, απέδειξαν ότι είναι ταχύτερη η εκτίμηση της συγκεντρωμένης συνάρτησης μέγιστης πιθανοφάνειας, για όλες τις τιμές της δ στο φάσμα τιμών της, και η εκτίμηση της βέλτιστης τιμής της, ως εκείνη που μεγιστοποιεί τη συνάρτηση στο φάσμα αυτό (αναφέρεται στον Elhorst, 2003).

<u>Βήμα 3</u>: Εν συνεχεία, εκτιμώνται διαδοχικά ξανά οι β και σ^2 από τις εξισώσεις (3.18) και (3.19).

- Η περίπτωση τυχαίων χωρικών επιδράσεων

Εδώ, όπως έχει αναφερθεί, οι μη παρατηρήσιμες χωρικές διαφοροποιήσεις είναι ασυσχέτιστες με τις ανεξάρτητες μεταβλητές του υποδείγματος. Για την εκτίμηση των παραμέτρων του υποδείγματος, β και δ, ακολουθείται η ίδια διαδικασία με αυτή των σταθερών επιδράσεων. Κατόπιν αυτού, οι αρχικές μεταβλητές $y_{i,t}$ και $x_{i,t}$ – της εξίσωσης (3.2.11) – μετασχηματίζονται ως εξής:

$$y^o_{i,t}(\varphi) = y_{i,t} - (1-\varphi)\frac{1}{T}\sum_{t=1}^{T} y_{i,t}$$

και (3.23)

$$x^o_{k,i,t}(\varphi) = x_{k,i,t} - (1-\varphi)\frac{1}{T}\sum_{t=1}^{T} x_{k,i,t}$$

όπου, σύμφωνα με τον Elhorst (2003), φ είναι το αποδιδόμενο βάρος στο διαστρωματικό τμήμα των δεδομένων, με την ιδιότητα: $0 < \varphi^2 = \sigma^2/(T\sigma_\mu^2 + \sigma^2) \leq 1$. Να σημειωθεί ότι για $\varphi = 0$, προφανώς, το υπόδειγμα τυχαίων χωρικών επιδράσεων ανάγεται σε υπόδειγμα σταθερών χωρικών επιδράσεων

Αντίστοιχα με την περίπτωση των σταθερών χωρικών επιδράσεων, το χωρικό υπόδειγμα του Durbin, της, εκφράζεται σε συνεπτυγμένη μορφή ως:

$$Y_t^o(\varphi) = \rho\, WY_t^o(\varphi) + Z_t^o(\varphi)\,\beta + e_t(\varphi) \qquad (3.24)$$

και ισοδύναμα:

$$e_t(\varphi) = (I_N - \rho W)Y_t^o(\varphi) - Z_t^o(\varphi)\,\beta \qquad (3.25)$$

όπου $Y_t^o = [y^o_{1,t}(\varphi), \ldots, y^o_{N,t}(\varphi)]'$, διαστάσεων $N\times 1$, $\beta = [\alpha, \beta_1, \ldots, \beta_k, \theta_1, \ldots, \theta_k]'$, διαστάσεων $(1+2k) \times 1$, $Z_t^o = [I_N', x^o_{1,i,t}(\varphi)', \ldots, x^o_{k,i,t}(\varphi)', Wx^o_{1,i,t}(\varphi)', \ldots, Wx^o_{k,i,t}(\varphi)']$, διαστάσεων $N \times (1+2k)$, και $e_t(\varphi) = [e_{1,t}(\varphi), \ldots, e_{N,t}(\varphi)]'$, διαστάσεων $N \times 1$, I_N' είναι το μοναδιαίο διάνυσμα $N \times 1$.

Η λογαριθμική συνάρτηση πιθανοφάνειας της εξίσωσης (3.25) είναι η εξής:

$$ln(L) = -\frac{NT}{2} ln(2\pi\sigma^2) + T\, ln|I_N - \rho W| + \frac{N}{2} ln(\varphi^2) - \frac{1}{2\sigma^2}\sum_{t=1}^{T} e_t'(\varphi) e_t(\varphi)$$

ή (3.26)

$$ln(L) = -\frac{NT}{2} ln(2\pi\sigma^2) + T\, ln|I_N - \rho W| + \frac{N}{2} ln(\varphi^2) - \frac{1}{2\sigma^2} \sum_{t=1}^{T} ((I_N - \rho W) Y_t^o(\varphi) - Z_t^o(\varphi)\beta)'((I_N - \rho W) Y_t^o(\varphi) - Z_t^o(\varphi)\beta))$$

Για δεδομένο φ, οι συνθήκες πρώτης τάξης – ως προς β και σ^2 – της εξίσωσης (3.26) καταλήγουν σε αντίστοιχους εκτιμητές των β και σ^2 με τις εξισώσεις (3.18)-(3.20) και (3.19)-(3.20), αντίστοιχα. Επομένως, δεδομένων των β, δ και σ^2, η εκτίμηση της παραμέτρου φ πραγματοποιείται εισάγωντας μια συγκεντρωμένη λογαριθμική συνάρτηση πιθανοφάνειας, της μορφής:

$$ln(L_c(\varphi)) = -\frac{NT}{2} + ln(e_t(\varphi)'e_t(\varphi)) + \frac{N}{2} ln(\varphi^2) \tag{3.27}$$

η οποία μπορεί να μεγιστοποιηθεί από τα κατάλοιπα των παλινδρομήσεων της Y^o στην W Y^o και Z^o, όπου $Y^o = [Y_1^o, ..., Y_T^o]'$ και $Z^o = [Z_1^o, ..., Z_T^o]'$, διαστάσεων $NT\times1$ και $NT\times(1+2k)$, αντίστοιχα. Ομοίως με προηγουμένως, θέτονται αρχικές τιμές για τις άγνωστες παραμέτρους και εφαρμόζεται η επαναληπτική διαδικασία που αναλύσαμε παραπάνω μέχρι τα αποτελέσματα να συγκλίνουν αριθμητικά (Elhorst, 2003).

3.2 Σχετικοί Έλεγχοι

3.2.1 Επιλογή υποδείγματος σταθερών ή τυχαίων επιδράσεων

Εφαρμόζοντας το Hausman test (Baltagi, 2005) πραγματοποιείται έλεγχος για την καταλληλότητα του μοντέλου τυχαίων επιδράσεων (REM) έναντι του μοντέλου σταθερών επιδράσεων (FEM). Το στατιστικό ελέγχου Hausman (Hausman, 1978) είναι το εξής:

$$\hat{h} = \frac{(\hat{\beta}_{FE} - \hat{\beta}_{RE})^2}{Var(\hat{\beta}_{FE}) - Var(\hat{\beta}_{RE})} \tag{3.28}$$

όπου $\hat{\beta}_{FE}$ and $\hat{\beta}_{RE}$ είναι $k\times1$ διανύσματα των συντελεστών εκτίμησης fixed και random effects αντιστοίχως, ενώ while $Var(\hat{\beta}_{FE})$ και $Var(\hat{\beta}_{RE})$ οι αντίστοιχοι

πίνακες διακύμανσης-συνδιακύμανσης *k×1*, και *k* το πλήθος των εκτιμώμενων παραμέτρων.

Το στατιστικό ελέγχου υπό τη μηδενική υπόθεση ακολουθεί ασυμπτωτικά την κατανομή X^2 με k βαθμούς ελευθερίας (ο αριθμός των ερμηνευτικών μεταβλητών του μοντέλου χωρίς τον σταθερό όρο). Η απόρριψη της μηδενικής υπόθεσης σημαίνει ότι το REM δεν είναι κατάλληλο για την εκτίμηση του μοντέλου και επιλέγουμε το FEM, ενώ η αποδοχή της ότι το REM είναι κατάλληλο για την εκτίμηση του μοντέλου.

Τέλος, να σημειωθεί ότι ο στατιστικός έλεγχος του Hausman μπορεί να χρησιμοποιηθεί ακόμα και στην περίπτωση που το μοντέλο επεκτείνεται και συμπεριλαμβάνει είτε χωρική αυτοσυσχέτιση λάθους είτε μια χωρικά εξαρτημένη μεταβλητή χρονικής υστέρησης αλλάζοντας τους βαθμούς ελευθερίας σε *k+1* καθώς το μοντέλο χωρικής έλλειψης έχει μία παραπάνω ερμηνευτική μεταβλητή.

3.2.2 Έλεγχος στασιμότητας (Stationarity test)

Πρωταρχικό βήμα είναι ο έλεγχος στασιμότητας των μεταβλητών, καθώς η μη ύπαρξη στασιμότητας είναι πιθανό να οδηγήσει σε μεροληπτικά αποτελέσματα. Εφαρμόστηκε ο έλεγχος Levin, Lin και Chu (2002):

Οι Levin, Lin και Chu ανέπτυξαν μια διαδικασία ελέγχου μοναδιαίας ρίζας για panel δεδομένα, που βασίζεται στον επαυξημένο έλεγχο των Dickey-Fuller (ADF)[21] ακολουθώντας την παρακάτω εξίσωση:

$$\Delta y_{i,t} = c_i + \rho_i y_{i,t-1} + \sum_{j=1}^{k} \theta_j \Delta y_{i,t-j} + \varepsilon_{it} \qquad (3.29)$$

όπου $t=1,\ldots,T$ είναι οι χρονικές περίοδοι, και $i=1,\ldots,N$ αριθμός των δια-στρωμάτων.

[21] βλ. ΠΑΡΑΡΤΗΜΑ ΙΙ: Επαυξημένος Έλεγχος Μοναδιαίας Ρίζας των Dickey-Fuller (ADF Test).

Βασική υπόθεση του ελέγχου είναι η ομοιογένεια (homogeneity) των συντελεστών σε ένα δυναμικό, αυτοπαλίνδρομο υπόδειγμα panel για όλα τα μέλη των στρωμάτων. Συγκεκριμένα, ο παραπάνω έλεγχος υποθέτει ότι κάθε στρώμα έχει το ίδιο αυτοπαλίνδρομο συντελεστή πρώτης τάξης ρ_i, δηλαδή $\rho_i=\rho$, αλλά επιτρέπει την επίδραση της χρονικής τάσης μεμονωμένα σε κάθε στρώμα. Οι χρονικές υστερήσεις της εξαρτημένης μεμεταβλητής μπορούν να προστεθούν στο υπόδειγμα για τον τμηματικό συσχετισμό στα λάθη.

Η μηδενική υπόθεση του ελέγχου είναι H_0: $\rho=\rho_i=0$, που εκφράζει την ύπαρξη μοναδιαίας ρίζας, και η εναλλακτική H_1: $\rho=\rho_i<0$, που σημαίνει ότι έχουμε υπόθεση στασιμότητας.

Για την υλοποίηση του ελέγχου γίνεται χρήση της συγκεντρωτικής t-στατιστικής του εκτιμητή που βασίζεται στο εκτιμώμενο ρ: $t_\rho = \frac{\hat{\rho}}{se(\hat{\rho})}$.[22]

3.2.3 Έλεγχος μη-αιτιότητας Granger (No-Causality test)

Προκειμένου να αντιμετωπιστεί το πρόβλημα της ενδογένειας εφαρμόζεται η Μέθοδος Μεγίστης Πιθανοφάνειας. Ένας τρόπος ελέγχου της ύπαρξης ή μη ενδογένειας των επεξεγηματικών μεταβλητών είναι ο έλεγχος ύπαρξης ή μη αντίστροφης αιτιότητας (reverse causality), δηλαδή από την εξαρτημένη μεταβλητή προς τις ανεξάρτητες[23].

---

[22] Η μηδενική υπόθεση έχει νόημα κάτω από συγκεκριμένες προϋποθέσεις, αλλά η εναλλακτική είναι πολύ ισχυρή για να πραγματοποιηθεί σε ενδιαφέρουσες εμπειρικές περιπτώσεις. Βασική προϋπόθεση για τον Levin-Lin-Chu (LLC) έλεγχο είναι η σχέση $\frac{\sqrt{N_T}}{T} \to 0$, ενώ επαρκείς συνθήκες είναι ότι $\frac{N_T}{T} \to 0$ και $\frac{N_T}{T} \to k$. (*N_T σημαίνει ότι η διαστρωματική διάσταση N είναι μονοτονική συνάρτηση σε σχέση με την διάσταση T του χρόνου). Ο συγκεκριμένος έλεγχος είναι πιο αξιόπιστος όταν το N παίρνει τιμές μεταξύ 10 και 250 και το T μεταξύ 5 και 250. Εάν το T είναι πολύ μικρό ο έλεγχος αποδυναμώνεται. Κύριο μειονέκτημα του ελέγχου είναι ότι βασίζεται στην υπόθεση διαστρωματικής εξάρτησης. Επίσης, η μηδενική υπόθεση ότι κάθε χρονολογική σειρά περιέχει μοναδιαία ρίζα είναι πολύ περιοριστική.

[23] Παρ' όλο που η ανάλυση παλινδρόμησης αποτυπώνει ουσιαστικά τη συσχέτιση μεταξύ δύο ή περισσότερων μεταβλητών, δεν συνεπάγεται αιτιότητα. Η «Αιτιότητα κατά Granger» (Granger Causality) μιας μεταβλητής x_t ως προς μια μεταβλητή y_t αναφέρεται στο κατά πόσο η πρόσφατη και προηγούμενη πληροφορία για τη x_t συμβάλλει στην πρόβλεψη και την καλύτερη ερμηνεία του

Με γνώμονα αυτό, πριν την εκτίμηση του υποδείγματος εφαρμόζουμε τον έλεγχο αντίστροφης μη-αιτιότητας (reverse no-causality) του Granger μέσω της μεθόδου Dumitrescu and Hurlin (2012).

Οι Dumitrescu and Hurlin (2012) βασιζόμενοι στη θεωρία του Granger πρότειναν μία επέκταση του ελέγχου αιτιότητας σε δεδομένα panel.

Το σχετικό υπόδειγμα είναι το εξής:

$$y_{i,t} = \alpha_i + \sum_{L=1}^{k} \gamma_{iL} y_{i,t-L} + \sum_{k=1}^{k} \beta_{iL} x_{i,t-L} + \varepsilon_{i,t} \text{ , } i=1,\dots, N \text{ και } t=1,\dots, T \qquad (3.30)$$

όπου $x_{i,t}$ και $y_{i,t}$ είναι οι μεμονωμένες παρατηρήσεις δύο στάσιμων μεταβλητών για τη διαστρωματική μονάδα i την περίοδο t. Οι συντελεστές μπορεί να διαφέρουν μεταξύ των διαστρωματικών μονάδων αλλά είναι χρονικά αμετάβλητοι (άρα πρόκειται για περίπτωση μοντέλου σταθερών επιδράσεων), ενώ η υστέρηση K είναι ίδια για όλες τις ομάδες (διαστρωματικές μονάδες). Επίσης, τα δεδομένα panel πρέπει να είναι ισορροπημένα (balaned).

Η μηδενική υπόθεση του ελέγχου είναι η εξής:

$$H_0 : \beta_{i1} = \cdots = \beta_{iK} = 0, \ \forall i = 1,\dots,N$$

και εκφράζει την απουσία αιτιότητας για όλα τα στοιχεία του panel. Η εναλλακτική υπόθεση είναι:

$$H_1 : \beta_{i1} = 0 \text{ ή } \cdots \text{ ή } \beta_{iK} = 0, \ \forall i = 1,\dots,N_1 \text{ με } N_1 < N$$

$$\text{και } \beta_{i1} \neq 0 \text{ ή } \cdots \text{ ή } \beta_{iK} \neq 0, \ \forall i = 1 + N_1,\dots,N$$

Στο πλαίσιο αυτό, οι Dumitrescu and Hurlin (2012) προτείνουν την ακόλουθη διαδικασία:

1. Εκτέλεση των N μεμονωμένων υποδειγμάτων που περιλαμβάνονται στη σχέση (3.30)
2. Εφαρμογή του ελέγχου F των K-σε πλήθος υποθέσεων ων $\beta_{i1} = \cdots = \beta_{iK} = 0$ προκειμένου να υπολογιστούν τα μεμονωμένα (για κάθε διαστρωματική μονάδα) W_i στατιστικά του ελέγχου Wald
3. Υπολογισμός του μέσου των στατιστικών ελέγχων W_i:
   a. $\overline{W} = \frac{1}{N} \sum_{i=1}^{N} W_i$ (3.31)

τρέχοντος y_t . Πρακτικά, η αιτιότητα ελέγχεται αν οι συντελεστές των με υστέρηση τιμών του x είναι στατιστικά σημαντικοί.

Να σημειωθεί ότι η απόρριψη της μηδενικής υπόθεσης δεν αποκλείει τη μη αιτιότητα για ορισμένες διαστρωματικές μονάδες.

Οι Dumitrescu and Hurlin (2012) απέδειξαν ότι το W συμπεριφέρεται ασυμπτωτικά καλά και μπορεί πραγματικά να χρησιμοποιηθεί για να διερευνήσει την αιτιότητα των panel δεδομένων.

Υπό την παραδοχή ότι τα στατιστικά στοιχεία Wald, W_i, είναι ανεξάρτητα και ταυτόσημα διανεμημένα σε όλες τις διαστρωματικές μονάδες μπορεί να αποδειχθεί ότι το τυποποιημένο στατιστικό ελέγχου $\bar{Z}$ όταν το T, σε πρώτη φάση, είναι αρκετά μεγάλο (τείνει στο άπειρο) και σε δεύτερη φάση το N είναι αρκετά μεγάλο (τείνει στο άπειρο), ακολουθεί τυπική κανονική κατανομή.

$$\bar{Z} = \sqrt{\frac{N}{2k}} \times (\bar{W} - k) \xrightarrow[T,N\to\infty]{d} N(0,1) \tag{3.32}$$

Επιπρόσθετα, για ένα σταθερό T με $T > 5 + 3K$ το κατά προσέγγιση τυποποιημένο στατιστικό ελέγχου $\tilde{Z}$ ακολουθεί τυπική κανονική κατανομή.

$$\tilde{Z} = \sqrt{\frac{N}{2k} \times \frac{T-3k-5}{T-2k-3}} \times \left(\frac{T-3k-3}{T-3k-1} \times \bar{W} - k\right) \xrightarrow[N\to\infty]{d} N(0,1) \tag{3.33}$$

Αν, τελικά, οι τιμές $\bar{Z}$, $\tilde{Z}$ είναι μεγαλύτερες από τις αντίστοιχες τυπικές κρίσιμες τιμές, τότε απορρίπτεται η H_0 και, συνεπώς, υπάρχει αιτιότητα του Granger.

3.2.4 Έλεγχος Ευστάθειας Συστήματος (Stability test)

Ο τελευταίος έλεγχος που θα πρέπει να πραγματοποιηθεί αφορά την ευστάθεια του υποδείγματος-συστήματος που έχουμε εκτιμήσει.

Σύμφωνα με τους Baltagi et al. (2018) και Fingleton and Szumilo (2019) για το Δυναμικό Υπόδειγμα Χωρικής Διάχυσης Durbin:

$$y_{it} = at + \tau y_{it-1} + \rho \sum_{j=1}^{N} w_{ij}\ y_{jt} + x_{it}\ \beta + \eta \sum_{j=1}^{N} w_{ij}\, y_{jt-1} +$$

$$+\theta \sum_{j=1}^{N} w_{ij}\ x_{ijt} + \mu_i + u_{it} \tag{3.34}$$

οι συνθήκες ευστάθειας πληρούνται όταν οι ιδιοτιμές της μήτρας $A_N = (B_N)^{-1} C_N$, όπου $B_N = I_N - \rho W_N$ και $C_N = \tau I_N + \eta W_N$ με ρ τον εκτιμητή του συντελεστή της χωρικής

διάχυσης της εξαρτημένης μεταβλητής, *τ* τον εκτιμητή του συντελεστή της χρονικής υστέρησης της εξαρτημένης μεταβλητής και *η* τον εκτιμητή του συντελεστή της χωρο-χρονικής υστέρησης της εξαρτημένης μεταβλητής, όταν είναι κατ' απόλυτη τιμή μικρότερες της μονάδας. Επομένως, αν η μέγιστη κατ' απόλυτη τιμή χαρακτηριστική ρίζα (ιδιοτιμή) της μήτρας $A_N = (B_N)^{-1} C_N$ είναι μικρότερη της μονάδας τότε το υπόδειγμα που εκτιμάται χαρακτηρίζεται ευσταθές (stable).

3.3 Μελέτη Περίπτωσης

3.3.1 Εξειδίκευση του υποδείγματος

Η συγκεκριμένη μελέτη επικεντρώνεται στις επιδράσεις ορισμένων αγωνιστικών και διοικητικών παραγόντων στο πλήθος νικών στο σύνολο των 82 αγώνων της κανονικής περιόδου μιας ομάδας του NBA.

Ακολουθώντας τις εργασίες των Stefani (1977a,b), Yang (2015) και Manner (2015), αφετηρία της μελέτης μας αποτελεί ένα απλό υπόδειγμα λογαριθμικής, γραμμικής παλινδρόμησης με εξαρτημένη μεταβλητή το πλήθος νικών στους 82 αγώνες της κανονικής περιόδου του NBA:

$$lnw_{82,\,it} = a + \beta_1 lnefgprc_{it} + \beta_2 lnassrt_{it} + \beta_3 lnorebprc_{it} + \\ + \beta_4 lndrebprc_{it} + \beta_5 lntovprc_{it} + \beta_6 lnpace_{it} + \\ \beta_7 lndpv_{it} + \beta_8 lnproll_{it} + \beta_9 lntrd_{it} \qquad (3.35)$$

Η χρήση του φυσικού λογαρίθμου είναι ένας τρόπος να μειωθούν οι επιδράσεις κλίμακας (scale effects) σε μεταβλητές που είναι εκφρασμένες με διαφορετικές μονάδες μέτρησης.

Ακολουθώντας τους Fingleton and Szumilo (2019), αρχικά, προκειμένου να λάβουμε υπόψη τις αλληλεξαρτήσεις μεταξύ των ομάδων προσθέτουμε στο υπόδειγμα χωρική διάχυση (spatial lag), τόσο των ανεξάρτητων μεταβλητών όσο και της εξαρτημένης:

$$lnw_{82,\,it} = \alpha + \rho \sum_{j=1}^{30} w_{ij}\, lnw_{82,jt} + \beta_1 lnefgprc_{it} + \beta_2 lnassrt_{it} + \beta_3 lnorebprc_{it} + \\ + \beta_4 lndrebprc_{it} + \beta_5 lntovprc_{it} + \beta_6 lnpace_{it} +$$

$$+\beta_7 lndpv_{it} + \beta_8 lnproll_{it} + \beta_9 lntrd_{it} + +\theta_7 \sum_{j=1}^{30} w_{ij}\, lndpv_{ijt} +$$
$$+\theta_8 \sum_{j=1}^{30} w_{ij}\, lnproll_{ijt} + \theta_9 \sum_{j=1}^{30} w_{ij}\, lntrd_{ijt}\ u_{it} \qquad (3.36)$$

Εν συνεχεία, προκειμένου να λάβουμε υπόψη και τη δυναμική συμπεριφορά της εξαρτημένης μεταβλητής εισάγουμε τον παράγοντα του χρόνου, t, στο υπόδειγμα (3.36) προσθέτοντας χρονική (temporal lag) αλλά και χωροχρονική (spatio-temporal lag)[24] υστέρηση της εξαρτημένης μεταβλητής. Συνεπώς, το παραπάνω υπόδειγμα ανάγεται σε ένα Δυναμικό Υπόδειγμα Χωρικής Διάχυσης Durbin, όπως αναλύθηκε στις ενότητες 3.1.1 και 3.1.2:

$$ln\, w_{82,\, it} = \tau ln w_{82,\, it-1} + \rho \sum_{j=1}^{30} w_{ij}\, ln w_{82,jt} + \beta_1 lnefgprc_{it} + \beta_2 lnassrt_{it} +$$
$$+\beta_3 lnorebprc_{it} + \beta_4 lndrebprc_{it} + \beta_5 lntovprc_{it} + \beta_6 lnpace_{it} +$$
$$+\ \beta_7 lndpv_{it} + \beta_8 lnproll_{it} + \beta_9 lntrd_{it} + \eta \sum_{j=1}^{30} w_{ij}\, ln w_{82,\, jt-1} +$$
$$+\theta_7 \sum_{j=1}^{30} w_{ij}\, lndpv_{ijt} + \theta_8 \sum_{j=1}^{30} w_{ij}\, lnproll_{ijt} + \theta_9 \sum_{j=1}^{30} w_{ij}\, lntrd_{ijt}\ u_{it} \qquad (3.37)$$

όπου η εξαρτημένη μεταβλητή lnw_{82} φυσικός λογάριθμος του w_{82} που εκφράζει το πλήθος νικών από τις συνολικά 82 αναμετρήσεις της κανονικής περιόδου, ενώ οι επεξηγηματικές αποτελούν τους φυσικούς λογαρίθμους τόσο των αγωνιστικών παραγόντων (*efgprc* = eFG%, *assrt* = Assist Ratio, *orebprc* = OREB%, *drebprc* = DREB%, tovprc = TOV%, *pace*) όσο και των μη αγωνιστικών που έχουν να κάνουν με τις διοικητικές αποφάσεις κάθε ομάδας και συγκεκριμένα με την πολιτική της κάθε ομάδας σχετικά τα draft pick της (*dpv* = draft pick value), τις ανταλλαγές παικτών (*trd* = trade) και τέλος η συνολική μισθοδοσία των ομάδων (*payroll*).

Στο υπόδειγμα (3.37) ο συντελεστής της εξαρτημένης μεταβλητής χωρικής υστέρησης ρ ονομάζεται συντελεστής χωρικής αυτοπαλινδρόμησης. Οι συντελεστές θ δείχνουν την ύπαρξη επιδράσεων διάχυσης των φυσικών λογαρίθμων των μη αγωνιστικών επεξηγηματικών μεταβλητών. Ενώ τα w_{ij} είναι τα στοιχεία μιας *N x N* θετικής, χωρικής μήτρας βαρών *W* που εκφράζει την αλληλεπίδραση μεταξύ των ομάδων *i* και *j*.

[24] Η χωροχρονική διάχυση μπορεί να ερμηνευτεί ως χωρική διάχυση που εξαπλώνεται με την πάροδο του χρόνου.

3.3.2 Χωρική Μήτρα βαρών

Ιδιαίτερη σημασία για τα χωρικά υποδείγματα έχει ο προσδιορισμός της χωρικής μήτρας βαρών *W*. Πρόκειται, όπως έχει αναφερθεί παραπάνω, για έναν θετικό, τετραγωνικό πίνακα *NxN* (στην προκειμένη περίπτωση *30x30*) ο οποίος αποτυπώνει την αλληλεξάρτηση μεταξύ των ομάδων του υποδείγματος.[25]

Τα στοιχεία της χωρικής μήτρας βαρών πρέπει να είναι μη στοχαστικά και εξωγενώς ορισμένα (Anselin, 1988). Επιπλέον, τα στοιχεία των γραμμών της *W* πρέπει να είναι κανονικοποιημένα προκειμένου να εισαχθούν στη χωρικό υπόδειγμα του Durbin. Αυτό σημαίνει, για παράδειγμα, να αθροίζουν στη μονάδα ($\sum_{j=1}^{N} w_{ij} = 1$). Μία μέθοδος κανονικοποίησης της μη κανονικοποιημένης μήτρας W_0, η οποία και εφαρμόστηκε στην παρούσα μελέτη, είναι να διαιρεθεί κάθε στοιχείο της, $w_{0,ij}$, με τη μεγαλύτερη χαρακτηριστική της ρίζα $\omega_{0,max}$, έτσι ώστε $w_{ij} = (1/\omega_{0,max})w_{0,ij}$ (Anselin, 1988, Elhorst, 2014) [26].

Ο προσδιορισμός της μήτρας βαρών στη συγκεκριμένη εργασία στηρίχτηκε στον τρόπο διεξαγωγής των αγωνιστικών στη λίγκα του NBA. Κάθε μία από τις 30 ομάδες της διοργάνωσης στα 82 παιχνίδια της κανονικής περιόδου αντιμετωπίζει:

- 4 φορές τις 4 ομάδες που ανήκουν στην ίδια «κατηγορία» (division)
- 4 φορές τις 6 από τις συνολικά 10 ομάδες των άλλων division της ίδιας περιφέρειας (conference) και 3 φορές τις υπόλοιπες 4
- 2 φορές τις 15 ομάδες της άλλης περιφέρειας (conference)

Ως εκ τούτου, στην κανονική περίοδο ο ανταγωνισμός είναι μεγαλύτερος μεταξύ των ομάδων της ίδιας κατηγορίας (division), αφού οι ομάδες παίζουν τα

[25] Στην προκειμένη περίπτωση, δεν έχουμε τόσο πολύ «χωρική αλληλεξάρτηση», αλλά οργανωσιακή, δεδομένου ότι η αλληλεξάρτηση των ομάδων εξαρτάται από το φορμάτ της λίγκας.

[26] Στο χωρικό υπόδειγμα του Durbin, η ευστάθεια του υποδείγματος απαιτεί την ισχύ της σχέσης $1/\omega_{min} < \rho < 1/\omega_{max}$, όπου ω_{min} και ω_{max} είναι η μικρότερη και η μεγαλύτερη χαρακτηριστική ρίζα της *W*, αντίστοιχα. Η κανονικοποίηση ανά γραμμή της *W*, εν μέρει προϋποθέτει ότι η τιμή του συντελεστή χωρικής αυτοσυσχέτισης ρ θα περιορίζεται στο διάστημα (-1, +1) (Elhorst, 2014).

περισσότερα παιχνίδια μεταξύ τους και είναι γεωγραφικά κοντά. Κατόπιν, μεγαλύτερος είναι ανταγωνισμός μεταξύ των ομάδων ίδιας περιφέρειας. Είναι λογικό ότι για κάθε ομάδα θα είναι δυσκολότερη η πορεία της κανονικής περιόδου προς τα playoffs αν μια ομάδα της ίδιας κατηγορίας ενισχυθεί, εφόσον την «συναντάει» περισσότερες φορές, παρά αν ενισχυθεί κάποια ομάδα μιας άλλης κατηγορίας και της άλλης περιφέρειας, τις οποίες «συναντάει» λιγότερες φορές. Βάση αυτού και του περιεχομένου της μελέτης, που αφορά την επίδραση στις νίκες της κανονικής περιόδου μιας ομάδας: (α) η αλληλεπίδραση μεταξύ των ομάδων διαφορετικών περιφερειών θεωρείται αμελητέα και ορίστηκε ίση με 0, (β) η αλληλεπίδραση μεταξύ ομάδων ίδιας περιφέρειας αλλά διαφορετικής κατηγορίας είναι μεγαλύτερη σε σχέση με την πρώτη και ορίστηκε ίση με 0.5 ενώ (γ) η αλληλεπίδραση μεταξύ ομάδων ίδιας κατηγορίας είναι ακόμα μεγαλύτερη και ορίστηκε ίση με 1.

Διαιρώντας κάθε στοιχείο της μήτρας W με τη μεγαλύτερη χαρακτηριστική της ρίζα (υπολογίστηκε η μέγιστη κατ' απόλυτη ιδιοτιμή ίση με 9) προέκυψε η κανονικοποιημένη μήτρα βαρών, η οποία και εφαρμόστηκε στο δυναμικό υπόδειγμα Durbin.[27]

[27] βλ. ΠΑΡΑΡΤΗΜΑ V: Μήτρα Βαρών (W) και Κανονικοποιημένη Μήτρα Βαρών (Wnorm).

Κεφάλαιο 4: Εφαρμογή

4.1 Δεδομένα και πηγές δεδομένων

Τα δεδομένα που χρησιμοποιούνται στην παρούσα έρευνα προέρχονται, κατά κύριο λόγο, από τη βάση δεδομένων της επίσημης ιστοσελίδας του NBA καθώς και άλλων σχετικών ιστοσελίδων όπως Basketball Reference, ESPN, Real GM και HOOPSHYPE.

Πιο αναλυτικά, τα αγωνιστικά δεδομένα της έρευνας αφορούν τις νίκες στο σύνολο των 82 παιχνιδιών της κανονικής περιόδου (w_{82}) και τα προηγμένα (advanced) στατιστικά: eFG%, AST Ratio, OREB%, DREB%, TOV% και PACE.[28] Πρόκειται για στατιστικά στοιχεία που καταγράφονται και υπολογίζονται για τις κανονικές περιόδους των αγωνιστικών περιόδων 2005-2006 έως 2019-2020. Τα στοιχεία προέρχονται από τον επίσημο ιστότοπο της λίγκας του NBA.[29]

Στην ιστορία του NBA, παρουσιάζονται μεταβολές στον τρόπο διεξαγωγής της διοργάνωσης, όσον αφορά τον τρόπο διαχωρισμού των ομάδων και επομένως τον καθορισμό των αγωνιστικών αναμετρήσεων. Επιπλέον, δεν είναι λίγες οι περιπτώσεις ομάδων που άλλαξαν όνομα ή και πόλη. Έως το 2004-2005 δεν αγωνίζονταν και οι 30 ομάδες που υπάρχουν τώρα στη λίγκα. Από εκείνη τη χρονιά και μετά εφαρμόστηκε το συγκεκριμένο φορμάτ που βλέπουμε τώρα με την είσοδο και της 30ης ομάδας στη διοργάνωση. Ως εκ τούτου, η περίοδος που μελετάμε δεν μπορεί να επεκταθεί νωρίτερα της season 2004-2005.

Τα δεδομένα μας, ωστόσο, δεν περιορίζονται μόνο στα αγωνιστικά. Λήφθηκαν υπόψη οι αποφάσεις της διοίκησης και του front office της εκάστοτε ομάδας σχετικά με το σχεδιασμό της για την επόμενη χρονιά.

---

[28] βλ. ΠΑΡΑΡΤΗΜΑ ΙΙΙ: Προηγμένα Στατιστικά (Advanced Stats).

[29] https://www.nba.com/stats/teams/advanced/?sort=W&dir=-1.

Καταρχάς, ο παράγοντας draft (Draft Pick Value - dpv) εκφράζει την απόφαση της κάθε ομάδας για το πώς επιθυμεί να διαχειριστεί την επιλογή της στο draft πριν την αγωνιστική περίοδο. Υπάρχουν ομάδες που επιλέγουν να στηριχτούν στον παίκτη που διαλέγουν και πραγματοποιούν τις κατάλληλες κινήσεις ώστε να «χτίσουν» μια ομάδα στην οποία ο συγκεκριμένος παίκτης θα έχει πρωταγωνιστικό ρόλο. Άλλες ομάδες επιλέγουν να ανταλλάξουν τον παίκτη την ίδια βραδιά του draft για κάποιον άλλο παίκτη που πιστεύουν ότι ταιριάζει καλύτερα στο ρόστερ τους. Επίσης, αν κάποια ομάδα επιθυμεί να γίνει διεκδικήτρια του τίτλου στο άμεσο μέλλον, συνήθως δεν επιλέγει να στηριχτεί στο draft της και για αυτό είτε το ανταλλάσσει στη συνέχεια, είτε το έχει συμπεριλάβει σε ανταλλαγή τις προηγούμενες χρονιές. Συνεπώς, η συγκεκριμένη μεταβλητή αντιπροσωπεύει την απόφαση που έλαβε κάθε ομάδα για την επιλογή της τη βραδιά του draft εκφρασμένη σε χρηματικές μονάδες (σε σταθερές τιμές του έτους 2022), μέσω των συμβολαίων που υπέγραψε (αν υπέγραψε) με πρωτοετή. Για τις επιλογές του πρώτου γύρου, όσο πιο μεγάλο είναι το συμβόλαιο, άρα και η τιμή της μεταβλητής draft, τόσο πιο «νωρίς» διάλεξε η ομάδα. Για τους παίκτες του δεύτερου γύρου, επειδή δεν υπάρχει κανονισμός για τα συμβόλαια και δεν είναι πολλοί οι παίκτες που αγωνίζονται στο NBA άμεσα, καθορίστηκε η τιμή της μεταβλητής ως το μικρότερο συμβόλαιο του NBA αντιστοίχως σε κάθε σεζόν. Τα δεδομένα πηγάζουν από τη Wikipedia και τις ιστοσελίδες Basketball Reference[30], ESPN[31], Real GM[32] και του NBA[33] για τα draft κάθε χρονιάς από το 2004-2005 έως 2018-2019, το ύψος συμβολαίων των rookies, τους συνολικούς μισθούς των ομάδων και το μικρότερο οικονομικά συμβόλαιο ανά σεζόν.

Εν συνεχεία, ο παράγοντας trade (trd), επίσης εκφρασμένος σε χρηματικές μονάδες (σε σταθερές τιμές του έτους 2022), εκφράζει το κατά πόσο μια ομάδα ευνοείται ή ρισκάρει από μια ανταλλαγή με βάση την αξία των εμπλεκόμενων παικτών και επιλογών draft. Μια ομάδα, προκειμένου να είναι άμεσα ανταγωνιστική, και ακόμα περισσότερο διεκδικήτρια του τίτλου, καταφεύγει σε παίκτες με μεγαλύτερη

[30] https://www.basketball-reference.com/draft.
[31] http://www.espn.com/nba/salaries.
[32] https://basketball.realgm.com.
[33] https://www.nba.com, https://pr.nba.com.

εμπειρία και ικανότητα, οι οποίοι κατά βάση αμοίβονται περισσότερο. Από την άλλη πλευρά, υπάρχουν ομάδες που αποφασίζουν να «χτιστούν» από την αρχή. Για τον σκοπό αυτό, επιλέγουν να στηριχτούν στο παρόν σε νεαρούς παίκτες ή draft picks, επενδύοντας, έτσι, στην προοπτική τους. Ουσιαστικά, ρισκάρουν επιλέγοντας να χάσουν κάποιους καλούς και έμπειρους παίκτες, με τους οποίους δεν είχαν τα επιθυμητά αγωνιστικά αποτελέσματα, και να επενδύσουν σε νεότερους, ελπίζοντας ότι μελλοντικά θα εξελιχθούν σε υψηλής αξίας παίκτες. Με γνώμονα αυτό, μελετήθηκαν οι ανταλλαγές μία προς μία για κάθε ομάδα ανά αγωνιστική περίοδο και έως τη λήξη των ελευθέρων μεταγραφών (free agency). Οι παίκτες ομαδοποιήθηκαν σε κατηγορίες και βαθμολογήθηκαν με αστέρια. Από την καλύτερη προς τη χειρότερη κατηγορία παικτών η βαθμολογία είναι 5, 4, 3, 2 και 1 αστέρι. Δεδομένων των συμβολαίων των παικτών, σε κάθε κατηγορία τα αστέρια «μετατρέπονται» σε χρηματικές μονάδες 50, 28, 13.5, 5 και 0.5 εκατομμυρίων δολαρίων αντιστοίχως. Συνεπώς, το trade υπολογίστηκε για κάθε ομάδα ανά σεζόν προσθέτοντας τη βαθμολογία των παικτών που πήρε και αφαιρώντας αυτών που έδωσε. Αν το trade είναι κοντά στο 0 σημαίνει η ομάδα δεν «τόλμησε» να κάνει ραγδαίες αλλαγές. Επίσης, θα δούμε τιμές του trade που είναι αρκετά μεγαλύτερες από το 0, γεγονός που σημαίνει ότι η ομάδα πιθανώς έχει στοχεύσει να φτάσει πολύ μακριά στο πρωτάθλημα. Αντιθέτως, τιμές του trade που είναι αρκετά μικρότερες από το 0 εκφράζει ένα ρίσκο που παίρνει η ομάδα και πιθανώς μια επιθυμία να αλλάξει ριζικά το ρόστερ της ανταλλάσσοντας καλούς παίκτες με «ρολίστες», μέτριους ή μελλοντικές draft επιλογές ώστε να θέσει νέα θεμέλια. Τα δεδομένα που λήφθηκαν από τις προαναφερθείσες ιστοσελίδες αφορούσαν μόνο τα συμβόλαια των παικτών ώστε να κατηγοριοποιηθούν ενώ από τη Wikipedia μελετήθηκαν όλες οι ανταλλαγές που έγιναν ανά αγωνιστική περίοδο για κάθε ομάδα.

Τέλος, ο παράγοντας payroll (proll) αφορά τη συνολική μισθοδοσία των παικτών, δηλαδή το ύψος των συμβολαίων που η διοίκηση υπογράφει με τους παίκτες. Τα δεδομένα είναι εκφρασμένα σε χρηματικές μονάδες (σε σταθερές τιμές του έτους 2022), όπως οι προηγούμενοι δυο παράγοντες, και πηγάζουν από τη βάση

δεδομένων των ιστοσελίδων HOOPSHYPE[34] και ESPN[35]. Ενώ για όλες τις ομάδες το salary cap κάθε season είναι ίδιο, ο τρόπος διαχείρισής του που επιλέγεται από ομάδα σε ομάδα διαφέρει. Υπάρχουν ομάδες που επιλέγουν να προσφέρουν περισσότερα συμβόλαια σε νεαρούς, ανερχόμενους, «ρολίστες» ή και απλά μέτριους παίκτες με απώτερο σκοπό είτε να στηριχτούν σε αυτούς για το μέλλον είτε για να κάνουν το λεγόμενο «tanking» και να αφήσουν «χώρο» στο salary cap για υψηλότερα συμβόλαια καλών παικτών τις επόμενες χρονιές. Άλλες ομάδες δίνουν ένα με δύο υψηλά συμβόλαια και τα υπόλοιπα μοιράζονται σε πρωτοετής, μέτριους και «ρολίστες» που υποστηρίζουν τους καλύτερους ποιοτικά παίκτες (stars). Υπάρχουν και ομάδες που επιθυμώντας να γίνουν πρωταθληματικές άμεσα υπογράφουν πολύ υψηλά συμβόλαια ξεπερνώντας το salary cap ή και πληρώνοντας luxury tax.

4.2 Εμπειρικά αποτελέσματα

Στη συγκεκριμένη παράγραφο θα παρουσιαστούν τα αποτελέσματα της εμπειρικής ανάλυσης από εφαρμογή της μεθοδολογίας που έχει προαναφερθεί στη βάση δεδομένων.

Αρχικά, πραγματοποιείται ο έλεγχος μοναδιαίας ρίζας Levin, Lin και Chu (2002) για τις μεταβλητές του υποδείγματος από τον οποίο προέκυψαν τα εξής:

Πίνακας 4.1: Αποτελέσματα LLC (Levin, Lin και Chu) Unit Root Test

Variable	lnw82	lnefgprc	lnassrt	lnorebprc	lndrebprc	lntovprc	lnpace	lndpv	lnproll	lntrd
p-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

[34] https://hoopshype.com/salaries
[35] http://www.espn.com/nba/salaries

Είναι εμφανές ότι η p-τιμή του ελέγχου είναι μικρότερη από το επίπεδο σημαντικότητας 0,05 και συνεπώς μπορούμε να απορρίψουμε τη μηδενική υπόθεση. Δηλαδή, όλες οι μεταβλητές μπορούν να χαρακτηριστούν ως στάσιμες στη χρονοσειρά των δεκαπέντε ετών της μελέτης.

Σειρά έχει ο έλεγχος μη ύπαρξης αιτιότητας Granger των Dumitrescu and Hurlin (2012) για δεδομένα panel. Στην παρούσα μελέτη, λόγω πεπερασμένης χρονοσειράς των δεδομένων, χρησιμοποιήσαμε το στατιστικό $\tilde{Z}$, το οποίο αναφέρεται στο μεθοδολογικό πλαίσιο, για την πραγματοποίηση του ελέγχου.

Πίνακας 4.2: Αποτελέσματα Dumitrescu and Hurlin (2012) Non-Causality Test

Optimal number of lags: AIC (lags tested: 1 to 3)	
H_0: dependent variable does not Granger-cause independent variable H_1: dependent variable does Granger-cause independent variable for at least one panelvar (id).	
Ind. Variable	**Z-bar tilde (p-value)**
Lnefgprc	0.878 (0.380)
Lnassrt	-0.629 (0.530)
Lnorebprc	0.282 (0.778)
Lndrebprc	1.533 (0.125)
Lntovprc	-0.333 (0.739)
Lnpace	0.387 (0.699)
Lndpv	1.153 (0.249)
Lnproll	1.347 (0.178)
Lntrd	0.470 (0.638)

Με τον έλεγχο μη αιτιότητας ελέγχουμε αν δεν υπάρχει «αντίστροφη αιτιότητα», από την εξαρτημένη προς τις ανεξάρτητες μεταβλητές. Οι μεγάλες τιμές της p-value δηλώνουν ότι δεχόμαστε τη μηδενική υπόθεση (H_0).

Ο τρίτος έλεγχος που κρίνεται απαραίτητος, βάση της μεθοδολογίας, είναι αυτός του Hausman, προκειμένου να αποφανθούμε αν το υπόδειγμα που θέλουμε να αναπτύξουμε είναι σταθερών ή τυχαίων επιδράσεων. Ωστόσο, λόγω περιορισμένης χρονοσειράς δεν επαρκούν οι βαθμοί ελευθερίας για να τρέξουν οι τυχαίες επιδράσεις (random effects). Ως εκ τούτου, προχωρήσαμε στην εφαρμογή ενός μοντέλου σταθερών επιδράσεων.

Στη συνέχεια, εφόσον τα αποτελέσματα των ελέγχων που προηγήθηκαν επέτρεψαν να προχωρήσουμε στην εφαρμογή του Δυναμικού Υποδείγματος Χωρικής Διάχυσης Durbin, παρουσιάζονται συγκεντρωμένες οι μεταβλητές που εξετάστηκαν:

Πίνακας 4.3: Περιγραφικά Στατιστικά

Variable	Obs	Mean	Std. Dev.	Min	Max
w82	450	40.78325	12.54024	8.69697	73
lw82	450	40.83217	12.55856	8.69697	73
efgprc	450	50.35956	2.166542	43.9	56.9
assrt	450	16.85289	1.177462	14.1	21.2
orebprc	450	29.14644	2.683323	21.6	36.1
drebprc	450	70.82378	2.142806	65.2	77.5
tovprc	450	15.01933	1.181051	11.9	19
pace	450	95.06153	3.701057	87.39	105.51
dpv	450	2931074	2235084	0	1.62e+07
proll	450	9.03e+07	2.13e+07	4.44e+07	1.68e+08
trd	450	-114669.6	1.41e+07	-5.82e+07	7.28e+07

Πίνακας 4.4: Περιγραφικά Στατιστικά με φυσικό λογάριθμο

Variable	Obs	Mean	Std. Dev.	Min	Max
lnw82	450	3.653145	0.3497273	2.162975	4.290459
lnefgprc	450	3.91827	0.0428463	3.781914	4.041295
lnassrt	450	2.822107	0.069476	2.646175	3.586293
lnorebprc	450	3.368053	0.0929782	3.072693	3.586293
lndrebprc	450	4.259739	0.0302019	4.177459	4.350278
lntovprc	450	2.706241	0.078935	2.476538	2.944439
lnpace	450	4.553775	0.0386579	4.470381	4.658806
lndpv	450	13.54475	4.111609	0	16.60333
lnproll	450	18.29197	0.2265536	17.60819	18.93915
lntrd	450	-.15774	14.73417	-17.87885	18.10297

Περνώντας στο «κυρίως θέμα» που αποτελεί και το αντικείμενο της παρούσας έρευνας, τα αποτελέσματα από την εφαρμογή του υποδείγματος για το οποίο έχει γίνει λόγος έως τώρα φαίνονται στον Πίνακα 4.5:

Πίνακας 4.5: Αποτελέσματα Δυναμικού Υποδείγματος Χωρικής Διάχυσης Durbin

R-sq: within = 0.6517			Number of obs = 450		
between = 0.6682			Number of groups = 30		
overall = 0.6546			Panel length = 15		
Mean of fixed-effects = -12.8367					
Log-likelihood = 119.8707					
Main		**Wx**	**Direct**	**Indirect**	**Total**
Variables	**Coef. (z-statistic)**				
lnw82		-0.3409787 (-2.53)			
lnlw82	0.110416 (3.04) ***	0.214746 (1.16)	0.1078706 (3.01)	0.1412009 (0.98)	0.2490715 (1.63)
lnefgprc	5.19049 (14.44) ***	-	5.210402 (14.61)	-1.343807 (-3.00)	3.866595 (7.17)
lnassrt	0.6105188 (3.42) ***	-	0.6216495 (3.37)	-0.1613088 (-2.16)	0.4603407 (3.18)
lnorebprc	1.422259 (10.28) ***	-	1.42892 (10.61)	-0.3694182 (-2.88)	1.059501 (6.72)
lndrebprc	2.662348 (6.56) ***	-	2.683578 (6.38)	-0.6916185 (-2.75)	1.99196 (4.99)
lntovprc	-1.225436 (-8.21) ***	-	-1.236443 (-8.09)	0.3196116 (2.83)	-0.9168314 (-5.92)
Lnpace	-1.829983 (-4.64) ***	-	-1.862055 (-4.64)	0.4805469 (2.53)	-1.381508 (-4.05)
Lndpv	-0.0060628 (-2.62) ***	-0.0059205 (-0.71)	-0.0061163 (-2.64)	-0.002956 (-0.44)	-0.0090723 (-1.30)
Lnproll	0.1205262 (1.64)	-0.6316177 (-5.90)	0.1346341 (1.78)	-0.5337181 (-4.83)	-0.3990841 (-4.67)
Lntrd	0.0012776 (2.02) *	-0.0023428 (-0.87)	0.0013589 (2.20)	-0.0022828 (-1.08)	-0.0009239 (-0.41)

Σε μια πρώτη ανάγνωση παρατηρούμε ότι ο συντελεστής προσδιορισμού R^2 είναι σχετικά καλός, αφού είναι πάνω από 0.65 και όχι σημαντικά μακριά από τη μονάδα, κάτι που εκφράζει έναν ικανοποιητικό βαθμό καταλληλότητας του υποδείγματος.

Ως προς τη στατιστική σημαντικότητα των ανεξάρτητων μεταβλητών, παρατηρούμε τα εξής: Από τις κύριες επιδράσεις μόνο η p-τιμή του φυσικού λογαρίθμου των συνολικών μισθών των ομάδων είναι μεγαλύτερη από το επίπεδο σημαντικότητας (0.05), γεγονός που σημαίνει ότι ο παράγοντας αυτός δεν είναι στατιστικά σημαντικός στο 5%. Ωστόσο, φαίνεται ότι είναι στατιστικά σημαντικός οριακά στο 10%. Για τις υπόλοιπες μεταβλητές η μηδενική υπόθεση του ελέγχου που εκφράζει τη σημαντικότητά τους γίνεται δεκτή, αφού η p-τιμή είναι μικρή. Όσον αφορά τις έμμεσες επιδράσεις (με τη μήτρα βαρών W), μόνο ο φυσικός λογάριθμος του payroll είναι στατιστικά σημαντικός σε επίπεδο σημαντικότητας 5%.

Αναφορικά με το πρόσημο της επίδρασης, με εξαίρεση τον παράγοντα του ποσοστού των επιθέσεων που καταλήγουν σε λάθος (TOV%) και το πλήθος κατοχών ανά αγώνα, οι αγωνιστικοί παράγοντες έχουν θετική επίδραση με το πλήθος νικών. Όπως είναι λογικό, όσο μεγαλύτερη η τιμή των δεικτών τόσο καλύτερη είναι η απόδοση της ομάδας επιθετικά και, συνεπώς, οδηγείται σε περισσότερα νικηφόρα αποτελέσματα.

Το ποσοστό επιθέσεων που καταλήγουν σε λάθος έχει αρνητική επίδραση στο πλήθος νικών, καθώς αγωνιστικά μια ομάδα στοχεύει στο να μειώσει όσο το δυνατόν περισσότερο τα λάθη της. Λιγότερες επιθέσεις που καταλήγουν σε λάθος συνεπάγονται περισσότερες ευκαιρίες για επιτυχημένες και εύστοχες επιθέσεις.

Το πλήθος κατοχών ανά 48 λεπτά παιχνιδιού (pace) επιδρά αρνητικά στο πλήθος νικών κατά την κανονική περίοδο. Το γεγονός αυτό εκφράζει ότι οι ομάδες με λιγότερες κατοχές (ίσως πιο στοχευμένες και εύστοχες επιθέσεις) είναι πιθανότερο να τερματίσουν με περισσότερες νίκες συγκριτικά με ομάδες που παίζουν με γρηγορότερο ρυθμό.

Συνεχίζοντας με μη αγωνιστικούς παράγοντες, η αξία του draft έχει αρνητική επίδραση στον αριθμό νικών. Λόγω της διαδικασίας επιλογής των παικτών στο draft, όσο μεγαλύτερη η αξία του pick τόσο υψηλότερα επέλεξε η ομάδα παίκτη και, συνεπώς, τόσο χαμηλότερα κατατάχθηκε την περασμένη αγωνιστική περίοδο.

Αντίθετα, οι παράγοντες της αξίας των ανταλλαγών και των συνολικών μισθών της ομάδας, όπως είναι αναμενόμενο, έχουν θετική επίδραση με το πλήθος νικών της

κανονικής περιόδου. Όσο υψηλότερη η αξία των ανταλλαγών, τόσο μεγαλώνει και η αξία του ρόστερ της ομάδας. Αντιστοίχως, όσο υψηλότεροι οι συνολικοί μισθοί μιας ομάδας θεωρητικά έχει ποιοτικότερους παίκτες στο δυναμικό της, καθώς οι καλοί παίκτες έχουν και υψηλότερα οικονομικά συμβόλαια. Ως εκ τούτου, υψηλές τιμές των παραγόντων αυτών συνεπάγονται και περισσότερες αναμενόμενες νίκες στο σύνολο των 82 αγώνων της κανονικής περιόδου.

Επιπρόσθετα, η συνολική μισθοδοσία μιας ομάδας φαίνεται ότι έχει στατιστικά σημαντική επίδραση στο πλήθος νικών των υπολοίπων ομάδων, κάτι που δεν ισχύει για την αξία του draft και των ανταλλαγών. Πρόκειται, όπως είναι λογικό, για αρνητική επίδραση αφού όσο υψηλότερο το payroll μιας ομάδας τόσο πιο δυνατό θεωρείται το αγωνιστικό δυναμικό της και επομένως αποτελεί ισχυρότερο αντίπαλο.

Όσον αφορά τον παράγοντα του ρεκόρ νίκης-ήττας της προηγούμενης κανονικής περιόδου μιας ομάδας, φαίνεται ότι είναι θετικός και στατιστικά σημαντικός για το πλήθος νικών της στην τρέχουσα κανονική περίοδο. Ωστόσο, δεν επιδρά ισχυρά στο ρεκόρ των αντιπάλων της, αλλά σε ένα επίπεδο σημαντικότητας 75%.

Τελευταία ενέργεια αποτελεί ο έλεγχος ευστάθειας του εκτιμώμενου μοντέλου με τον τρόπο που αναπτύχθηκε στην τελευταία παράγραφο του τρίτου κεφαλαίου. Η μέγιστη (κατ' απόλυτη τιμή) ιδιοτιμή του πίνακα $A_N = (B_N)^{-1} C_N$, όπου $B_N = I_N - \rho W_N$ και $C_N = \tau I_N + \eta W_N$, σύμφωνα με τα παραπάνω αποτελέσματα υπολογίστηκε ίση με 0,2466. Εφόσον πρόκειται για τιμή μικρότερη της μονάδας το υπόδειγμα ικανοποιεί την υπόθεση της ευστάθειας.

Κεφάλαιο 5: Συμπεράσματα

5.1 Συμπεράσματα

Στο κεφάλαιο αυτό αναλύονται τα αποτελέσματα από την εκτίμηση του Δυναμικού Υποδείγματος Διάχυσης Durbin που επεξηγεί το πλήθος νικών μιας ομάδας μπάσκετ του NBA στα 82 παιχνίδια της κανονικής αγωνιστικής περιόδου συναρτήσει ορισμένων μεταβλητών.

Παράγοντες που επηρεάζουν άμεσα το πλήθος νικών μιας ομάδας αποτελούν τα αγωνιστικά στοιχεία: Αποτελεσματικότητα εντός πεδιάς (eFG%), Ποσοστό επιθέσεων που καταλήγουν σε ασσίστ (Assist Ratio), Ποσοστό κερδισμένων επιθετικών και αμυντικών ριμπάουντ από τα συνολικά διεκδικούμενα (Offensive & Defensive Rebounds%), Ποσοστό επιθέσεων που καταλήγουν σε λάθος (Turnovers%) και Πλήθος κατοχών ανά παιχνίδι (Pace) καθώς και τα μη αγωνιστικά: Αξία draft pick (Draft Pick Value) και ανταλλαγών (Trade).

Πιο συγκεκριμένα, ξεκινώντας από τον δείκτη eFG%, όσο μεγαλύτερος είναι τόσο καλύτερο είναι το ρεκόρ νίκης-ήττας μιας ομάδας. Ομάδες με μεγαλύτερη τιμή του δείκτη αυτού φαίνεται να έχουν περισσότερες νίκες σε σχέση με άλλες. Ο δείκτης assist ratio, από την άλλη, επιδρά στο πλήθος νικών λιγότερο σε σχέση με το eFG%, αφού μια ομάδα με μεγαλύτερο ποσοστό επιθέσεων που καταλήγουν σε ασσίστ έχει ελάχιστα καλύτερο ρεκόρ από μια άλλη ομάδα με μικρότερο ποσοστό.

Ομοίως, η θετική επίδραση του ποσοστού των επιθετικών και αμυντικών rebounds% με τις νίκες μιας ομάδας είναι επίσης αρκετά μικρότερη από αυτή του δείκτη eFG%. Μάλιστα, παρατηρείται ότι τα αμυντικά rebounds επιδρούν άμεσα στις νίκες σε μεγαλύτερο βαθμό συγκριτικά με τα επιθετικά- με σχεδόν διπλάσια επίδραση. Αυτό, πιθανώς, οφείλεται στο γεγονός ότι τα αμυντικά rebounds δίνουν μια νέα ολόκληρη κατοχή 24 δευτερολέπτων. Επίσης, κάθε αμυντικό rebound σημαίνει μια επιτυχημένη άμυνα αλλά κάθε επιθετικό δεν συνεπάγεται πάντα μια επιτυχημένη επίθεση. Γι’ αυτό, άλλωστε, οι ομάδες επενδύουν περισσότερο στο αμυντικό rebound, ενώ υπάρχουν κάποιες που επιλέγουν να μη διεκδικήσουν καθόλου επιθετικά rebounds.

Το ποσοστό επιθέσεων που οδηγούν σε λάθη έχουν αρνητική επίδραση στις νίκες, καθώς τα περισσότερα λάθη αποτελούν αρνητικό παράγοντα για το επιθυμητό αποτέλεσμα της νίκης. Όπως είναι λογικό, άλλωστε, όλες οι ομάδες στοχεύουν στον όσο δυνατό μικρότερο αριθμό λαθών.

Τελευταίο αγωνιστικό κριτήριο είναι το πλήθος κατοχών ανά 48 λεπτά παιχνιδιού που εκφράζει και τον ρυθμό του παιχνιδιού μιας ομάδας. Είναι εμφανής η στατιστική σημαντικότητα του δείκτη αλλά αξιοσημείωτη είναι η αρνητική επίδρασή του, δηλαδή όσο περισσότερες κατοχές έχει μια ομάδα σε έναν αγώνα τόσο λιγότερο πιθανό είναι να οδηγηθεί στη νίκη. Αυτό εξηγείται, πιθανώς, από το γεγονός ότι ομάδες με περισσότερες κατοχές έχουν ταχύτερο αγωνιστικό ρυθμό, κάτι που δεν είναι πάντα θετικό. Μια πιθανή ερμηνεία είναι ότι οι ομάδες αυτές καταλήγουν σε περισσότερα λάθη ή βιαστικές και όχι επιτυχημένες επιθέσεις. Δεν είναι τυχαίο ότι οι «χειρότερες» ομάδες συχνά επιλέγουν να εκτελούν γρήγορες επιθέσεις, εφόσον δε διαθέτουν την ποιότητα να στοχεύουν σε αργές και οργανωμένες επιθέσεις.

Συνεχίζοντας με τα μη αγωνιστικά κριτήρια, παρατηρείται ότι η αξία του draft pick αποτελεί τον δεύτερο παράγοντα με όχι τόσο ισχυρή άμεση επίδραση μετά το trade. Μια ομάδα με υψηλή αξία draft pick δεν έχει την ποιότητα να γίνει άμεσα ανταγωνιστική. Γι’ αυτό άλλωστε της δίνεται η ευκαιρία να επιλέξει νωρίς στο draft, προκειμένου να στοχεύσει σε έναν καλό, ανερχόμενο παίκτη.

Επίσης, παρατηρείται χαμηλή επίδραση του παράγοντα της αξίας των ανταλλαγών. Όσο πιο υψηλή η αξία της ανταλλαγής τόσο πιο ανταγωνιστική γίνεται η ομάδα αναβαθμίζοντας την αγωνιστική της ποιότητα άμεσα. Ωστόσο, η χαμηλή επίδραση οφείλεται ίσως στους κανονισμούς που καθορίζουν τις ανταλλαγές, προκειμένου να μη δίνεται πλεονέκτημα σε ομάδες να μετατρέπονται σε «υπερ-ομάδες» εις βάρος άλλων.

Τέλος, όσον αφορά τον παράγοντα των συνολικών μισθών έχει χαμηλή επιδραστικότητα στο ρεκόρ μιας ομάδας. Προφανώς, υψηλή μισθοδοσία συνεπάγεται καλύτερους ποιοτικά παίκτες και επομένως περισσότερες νίκες. Ωστόσο, η διαφορά στις νίκες σε σχέση με άλλες ομάδες με μικρότερη μισθοδοσία

δεν είναι σημαντική, διότι η συλλογική σύμβαση της διοργάνωσης δεν επιτρέπει στις ομάδες να ξεφύγουν οικονομικά με σχετικούς κανονισμούς.

Οι παράγοντες της αξίας του draft (Draft Pick Value), των συνολικών μισθών (Payroll) και των ανταλλαγών (Trade) επηρεάζουν και έμμεσα το πλήθος νικών μιας ομάδας στην κανονική περίοδο. Εφόσον πρόκειται για μια κλειστή λίγκα, οι αποφάσεις που λαμβάνει μια ομάδα επηρεάζουν αγωνιστικά και τις υπόλοιπες ομάδες.

Αναλυτικότερα, φαίνεται ότι μόνο το ύψος των μισθών μιας ομάδας επηρεάζει το πλήθος νικών των υπολοίπων ομάδων και μάλιστα περισσότερο απ' ότι επηρεάζει το δικό της. Συγκεκριμένα, όσο υψηλότερο είναι το payroll μιας ομάδας τόσο πιο ανταγωνιστική την καθιστά, αφού συνήθως οι καλύτεροι ποιοτικά παίκτες έχουν μεγαλύτερα συμβόλαια. Συνεπώς, ένας πιο δυνατός αντίπαλος συνεπάγεται μικρότερη πιθανότητα νίκης των υπολοίπων ομάδων απέναντί του και, ως εκ τούτου, μειώνεται σε μικρό βαθμό το πλήθος νικών.

Επιπλέον, τόσο η επίδραση της αξίας του draft pick όσο και των ανταλλαγών που πραγματοποιεί μια ομάδα είναι αρνητική για τους αντιπάλους της. Η επιλογή ενός καλού παίκτη στο draft αλλά και μία καλή ανταλλαγή καθιστούν μια ομάδα αμέσως ανταγωνιστική στη διεκδίκηση του πρωταθλήματος και δυσκολότερο αντίπαλο για τις υπόλοιπες ομάδες.

Ένας παράγοντας που σχετίζεται με την αγωνιστική εικόνα των ομάδων αποτελεί το ρεκόρ της προηγούμενης κανονικής περιόδου, το οποίο επηρεάζει άμεσα το ρεκόρ της επόμενης σεζόν με όχι ιδιαίτερα υψηλή επίδραση. Κατά κανόνα, για τις περισσότερες ομάδες οι μεταβολές των νικών από χρονιά σε χρονιά δεν είναι μεγάλες. Αυτός ίσως είναι και ο λόγος που ο παράγοντας των νικών της περασμένης περιόδου είναι στατιστικά σημαντικός. Παρατηρούμε, όμως, ότι σε βάθος ορισμένων ετών οι στρατηγικές επιλογές μιας ομάδας μπορούν να την καθιστούν πολύ καλύτερη ή χειρότερη. Επομένως, η χαμηλή επίδραση με το ρεκόρ της επόμενης σεζόν πιθανώς οφείλεται στο γεγονός ότι στη χρονοσειρά υπάρχουν ομάδες που είναι κατά περιόδους καλές ή κακές. Αυτή η κατάσταση, όμως, δεν

είναι ευσταθής, εξ' ου και οι απότομες μεταπτώσεις του πλήθους νικών που παρατηρείται μετά το πέρας κάποιου χρονικού διαστήματος.

Παρατηρείται ότι το ρεκόρ μιας ομάδας επηρεάζει περισσότερο τις υπόλοιπες σε σχέση με την ίδια, ωστόσο δεν είναι το ίδιο στατιστικά σημαντικό. Η καλαθοσφαίριση είναι ένα ευθέως ανταγωνιστικό άθλημα και δεν υπάρχει άλλο αποτέλεσμα πέρα από νίκη ή ήττα. Ακόμα, το πρωτάθλημα του NBA είναι «κλειστό». Οπότε, αν μια ομάδα κερδίζει ή χάνει περισσότερους αγώνες σε σχέση με την προηγούμενη χρονιά, οι αντίπαλοί της θα επηρεαστούν από τα αποτελέσματά της. Επειδή λοιπόν η επιτυχία μιας ομάδας είναι άμεσα συνυφασμένη με τα αποτελέσματά της, το ρεκόρ της προηγούμενης αγωνιστικής περιόδου καθορίζει και τις μελλοντικές της αποφάσεις καθώς και των ανταγωνιστών της. Ενδεικτικά, μια ομάδα μετά από επιτυχημένη χρονιά είναι πιθανό να διατηρήσει στο δυναμικό της τους καλούς της παίκτες και να προσελκύσει κι άλλους καλούς παίκτες. Έτσι, η ίδια θα γίνει ακόμα καλύτερη και παράλληλα δε θα επιτρέψει στους ανταγωνιστές της να ενισχυθούν. Από την άλλη, μια ομάδα που δεν πετυχαίνει τους στόχους που έχει θέσει είναι πιθανό να επιλέξει να γίνει χειρότερη παραχωρώντας τους καλύτερους της παίκτες. Έτσι, και η ίδια θα γίνει ευκολότερος αντίπαλος και οι ανταγωνιστές της θα ενισχυθούν.

Η έμμεση θετική επίδραση πιθανώς εξηγείται από το γεγονός ότι καμία ομάδα δεν έχει ολοκληρώσει την κανονική περίοδο αήττητη (ο Μ.Ο. των ηττών της ομάδας με το καλύτερο ρεκόρ στην κανονική περίοδο στη χρονοσειρά που μελετάμε είναι 17). Ουσιαστικά, οι ομάδες διακρίνονται σε αυτές που διεκδικούν το πρωτάθλημα, που ανταγωνίζονται για την πρόκριση στα play offs και που έχουν μπει ή πρόκειται να μπουν σε διαδικασία «ανοικοδόμησης». Βάση των αποτελεσμάτων, το προηγούμενο ρεκόρ κάθε ομάδας, ανεξάρτητα από την κατηγορία στην οποία ανήκει, επιδρά στις υπόλοιπες θετικά. Πιθανώς, αυτό οφείλεται στο ότι αν μία ομάδα κερδίζει περισσότερο, οι αντίπαλοί της προσπαθούν να ενισχυθούν για να την ακολουθήσουν. Ακόμα και για αυτούς που δεν είναι καλοί, έχουν την ευκαιρία να βελτιωθούν μέσω του draft. Αν μια ομάδα κερδίζει λιγότερο, προφανώς, οι αντίπαλοί της επωφελούνται από τη χαμηλότερη δυναμικότητά της. Κατά κύριο λόγο, λοιπόν, η έμμεση θετική επίδραση του προηγούμενου ρεκόρ της κανονικής

περιόδου εξηγείται από την ανταγωνιστικότητα και την οργανωσιακή διάρθρωση του πρωταθλήματος.

Τα Direct αποτελέσματα σχετίζονται με τις άμεσες επιδράσεις και οι παρατηρήσεις δε διαφέρουν σε σχέση με αυτά που αναφέρθηκαν παραπάνω. Από το πεδίο των Indirect, ομοίως, λαμβάνουμε αντίστοιχα συμπεράσματα για την επιδραστικότητα των παραγόντων draft pick, trade και payroll.

Αξίζει, ωστόσο, να σχολιάσουμε την έμμεση επιδραστικότητα των αγωνιστικών κριτηρίων. Φαίνεται πως αυτά που επηρεάζουν μία ομάδα θετικά επιδρούν αρνητικά στους αντιπάλους της. Ακόμα, οι δείκτες αυτοί διαθέτουν μια στατιστική σημαντικότητα και είναι εμφανές ότι επηρεάζουν περισσότερο την ίδια την ομάδα και λιγότερο τον αντίπαλο. Γενικά, λοιπόν, βελτίωση των αγωνιστικών παραγόντων συνεπάγεται μεν ότι γίνεται καλύτερη σε σχέση με τον εαυτό της και γίνεται δε πιο ανταγωνιστική ως προς τις αντιπάλους της.

Βιβλιογραφία

Baghal, T.A. (2012). “Are the “Four Factors” Indicators of One Factor? An Application of Structural Equation Modeling Methodology to NBA Data in Prediction of Winning Percentage”. *Journal of Quantitative Analysis in Sports*, Vol. 8(1), Manuscript 1355.

Baltagi, B.H. (2005). *Econometric Analysis of Panel Data*. England: John Wiley & Sons Ltd, The Atrium

Baltagi, B.H., Fingleton, B. and Pirotte, A. (2019). “A time-space dynamic panel data model with spatial moving average errors”, *Regional Science and Urban Economics,* Vol. 76, pp. 13–31.

Berri, D.J., Schmidt, M.B. and Brook, S.L. (2006). *The Wages of Wins: Taking Measure of the Many Myths in Modern Sport*. Stanford, CA: Stanford University Press

Bouayad, A. S., Le Gallo, J. and Védrine, L. (2018). *Handbook of Spatial Analysis: Theory and Application with R*, Insee Methods No 131, Eurostat.

Carlin, B. P. (1996) “Improved NCAA basketball tournament modeling via point spread and team strength information”, *The American* Statistician, Vol. 50, pp. 39–43.

Cohen, J. (2007). “Economic Benefits of Investments in Transport Infrastructure”, *OECD, International Transport Forum, OECD/ITF Joint Transport Research Centre* (Discussion Papers).

Doolittle, B. and Pelton, K. (2010) *Pro Basketball Prospectus.* United States (California): CreateSpace Independent Publishing Platform.

Elhorst, J.P. (2010). “Spatial panel data models”, In: *Handbook of Applied Spatial Analysis* (pp.377-407).

Elhorst, J.P. (2011) “Dynamic spatial panels: models, methods, and inferences”, *Journal of Geographical Systems*, Vol. 14, pp.5–28.

Elo, A. E. (1978). *The rating of chess players past and present.* New York: Arco.

Entine, O. A. and D. S. Small (2008). “The role of rest in the NBA home-court advantage”, *Journal of Quantitative Analysis in Sports,* Vol. 4(2), Article: 6.

Fahrmeir, L. and G. Tutz (1994). “Dynamic stochastic models for time-dependent ordered paired comparison systems”, *Journal of the American Statistical Association*, Vol. 89, pp. 1438–1449.

Fingleton, B. and Szumilo, N. (2019). “Simulating the impact of transport infrastructure investment on wages: A dynamic spatial panel model approach”, *Regional Science and Urban Economics*, Vol. 75, pp. 148-164

Glickman, M. E. (1993) *Paired comparison models with time-varying parameters*. Department of Statistics, Harvard University, Cambridge (Phd dissertation).

Glickman, M. E. (1999). “Parameter estimation in large dynamic paired comparison experiments”, *Applied Statistics*, Vol. 48, pp. 377–394.

Glickman, M. E. (2001). “Dynamic paired comparison models with stochastic variances”, *Journal of Applied Statistics,* Vol. 28, pp. 673–689*.*

Glickman, M. E. and H. S. Stern (1998). “A state-space model for national football league scores”, *Journal of the American Statistical Association*, Vol. 93, pp. 25–35.

Harville, D. A. (2003). “The selection of seeding of college basketball or football teams for postseason competition”, *Journal of the American Statistical Association*, Vol. 98, pp. 17–27.

Harville, D. A. and M. H. Smith (1994). “The home-court advantage: How large is it and does it vary from team to team”, *The American Statistician*, Vol. 48, pp. 22–29*.*

Hoffman, L. and Joseph, M. (2003). *A multivariate statistical analysis of the NBA*. University of Wisconsin River Falls (River Falls, WI) and Kentucky State University Frankfurt, KY

Hollinger, J. (2002). *Pro Basketball: Prospectus.* United States: Brassey’s

Hollinger, J. (2003). *Pro Basketball: Prospectus.* United States: University of Nebraska Press

Hollinger, J. (2004). *Pro Basketball: Forecast.* United States: Potomac Books

Hollinger, J. (2005). *Pro Basketball: Forecast.* United States: Potomac Books Inc.

Zhang, J., Piatt, D.M., Ostroff, D.H. and Wright, J.W. (2005). “Importance of In-Game Entertainment Amenities at Professional Sporting Events: A Case for NBA Season Ticket Holders”, *Journal of Contemporary Athletics*, Vol. 2, pp. 1-24.

Javier, G., Ibáñez, S.J., Santos, R.M.D. Nuno, L. and Sampaio, J. (2013). “Identifying basketball performance indicators in regular season and playoff games”, *Journal of Human Kinetics*, Vol. 36(1), pp. 161-168.

Jones, E.S. (2016) *PREDICTING OUTCOMES OF NBA BASKETBALL GAMES*. North Dakota State University (In Partial Fulfillment of the Requirements for the Degree of MASTER OF SCIENCE).

Jones, M.B. (2007). "Home advantage in the NBA as a game-long process", *Journal of Quantitative Analysis in Sports*, Vol. 3 (4), Article: 2.

Jones, M.B. (2008). "A note on team-specific home advantage in the NBA", *Journal of Quantitative Analysis in Sports,* Vol. 4 (3), Article: 5.

Kline, R.B. (2005). *Principles and Practices of Structural Equation Modeling*, *2nd Edition*, Guilford Press, New York, NY.

Knorr-Held, L. (2000). "Dynamic ratings of sports teams", *The Statistician,* Vol. 49, pp. 261–276.

Koopman, S.J. and Lit, R. (2015). "A dynamic bivariate poisson model for analyzing and forecasting match results in the english premier league", *Journal of the Royal Statistical Society A,* Vol. 178, pp. 167–186.

Kubatko, J., Oliver, D., Pelton, K. and Rosenbaum, D.T. (2007). "A Starting Point for Analyzing Basketball Statistics", *Journal of Quantitative Analysis in Sports, Vol. 3, Iss. 3, Article 1.*

Küpfer, E. (2005). Team Similarity. *APBRMetrics Forum*

Levin, A., Lin, C. and Chu, C. J. (2002). "Unit root tests in panel data: asymptotic and finite-sample properties", *Journal of Econometrics,* Vol. 108, pp. 1–24.

Lopez, L. and Webew, S. (2017). "Testing for Granger causality in panel data", *The Stata Journal,* Vol. 17, Number 4, pp. 972–984.

Lewis, M. (2003). *Moneyball: The Art of Winning an Unfair Game*. United States: W.W. Norton & Company.

Manner, H. (2015). *Modeling and forecasting the outcomes of NBA Basketball games.* Institute of Econometrics and Statistics, University of Cologne.

Μαρίνος, Θ. (2020). *Οι οικονομικές επιπτώσεις των επενδύσεων σε υποδομές μεταφορών στην Ελλάδα (1977-2015).* Αθήνα, Εθνικό Μετσόβιο Πολυτεχνείο (Διδακτορική Διατριβή).

Minghui, Q., Ridong, H. and Jianwei, C. (2015). “Estimation Of Spatial Dynamic Nonparametric Durbin Models With Fixed Effects”, *Eurasia Journal of Mathematics, Science & Technology Education*, Vol. 12(4), pp. 807-819.

Oliver, D. (2002). *Basketball on Paper: Rules and Tools for Performance Analysis.* United States (Washington D.C.): Brassey's, Inc.

Schwertman, N. C., McCready, T.A. and Howard, L. (1991). “Probability models for the NCAA regional basketball tournaments”, *The American Statistician,* Vol. 45, pp. 35–38.

Semykina, A. and Wooldridge, J.M. (2008). “Estimating Panel Data Models in the Presence of Endogeneity and Selection”, *Journal of Econometrics,* Vol. 157, Iss. 2, pp. 375-380.

Stefani, R. T. (1977a). “Football and basketball prediction using least squares”, *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-7, pp. 117–121.

Stefani, R. T. (1977b). “Improved least squares football, basketball, and soccer predictions”, *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-7, pp. 117–121.

Stern, H. S. (1994). “A brownian motion model for the progress of sports scores”, *Journal of the American Statistical Association*, Vol. 89, pp. 1128–1134.

Teramoto, M. and Cross, C.L. (2010). “Relative Importance of Performance Factors in Winning NBA Games in Regular Season versus Playoffs”, *Journal of Quantitative Analysis in Sports*, Vol. 6, Iss. 3, Article: 2.

Torres, R. A. (2013). *Prediction of NBA games based on machine learning methods.* University of Winconsin-Madison

Uudmae, J. (2017). *Predicting NBA game outcomes.* Leland Stanford Junior University

Wooldridge, J.M. (2003, 2012). *Introductory Econometrics: A Modern Approach, 2nd and 5th edition.* United States: South Western Publishing Co.

Yuanhao (Stanley), Y. (2015). *Predicting Regular Season Results of NBA Teams Based on Regression Analysis of Common Basketball Statistics*. University of California at Berkeley (A thesis submitted in fulfillment of the requirement for the degree of honors in Statistics).

Zhang, X. (2019). *Modeling of NBA Game Data and their Correlation Structure*. University of Houston (Degree of Doctor of Philosphy in Mathematics).

Διαδικτυακοί τόποι και πηγές δεδομένων

https://www.insee.fr/en/information/3635545

https://www.researchgate.net/publication

http://www.units.miamioh.edu/sumsri/sumj/2003/NBAstats.pdf

www.elsevier.com/locate/regec

www.sciencedirect.com

https://www.nba.com/stats/teams/advanced/?sort=W&dir=-1

https://www.basketball-reference.com/draft,http://www.espn.com/nba/salaries

https://basketball.realgm.com,https://basketball.realgm.com/nba/info/rookie_scale

https://basketball.realgm.com/nba/playoffs/history?fbclid=IwAR18QdyOjRtxxgWnSTaHikOIapbj7hmlNmv2-Tjgqvm-GluUzoF2rrorlaI

https://www.nba.com

https://pr.nba.com

https://hoopshype.com/salaries

http://www.espn.com/nba/salaries

https://www.nba.com/stats/help/glossary

https://www.landofbasketball.com/championships/year_by_year.htm

https://nbamath.com/adjusted-team-ratings/

https://www.forbes.com/sites/kurtbadenhausen/2019/02/06/nba-team-values-2019-knicks-on-top-at-4-billion/#283d644ae667

https://www.forbes.com/nba-valuations/list/

https://bleacherreport.com/articles/2733544-ranking-every-nba-franchise-on-historical-success

https://bleacherreport.com/articles/824669-ranking-the-most-successful-franchises-in-nba-history

https://bleacherreport.com/articles/1333672-big-market-team-nba-power-rankings

https://bleacherreport.com/articles/1325290-small-market-team-nba-power-rankings

https://bleacherreport.com/articles/870390-ranking-the-front-office-from-every-nba-team-from-worst-to-first

https://bleacherreport.com/articles/2787678-updated-nba-front-office-power-rankings-post-free-agency

https://www.reddit.com/r/nba/comments/axu1os/rank_the_front_offices/

https://www.stationindex.com/tv/tv-markets

https://sportsshow.net/most-popular-sports-in-america/

https://blog.statscore.com/what-are-the-5-most-popular-sports-in-the-usa/

https://www.totalsportek.com/worlds-popular-sports-eleven/

https://www.forbes.com/athletes/

https://www.statista.com/statistics/193467/total-league-revenue-of-the-nba-since-2005/

https://cbabreakdown.com/salary-cap-overview

https://cbabreakdown.com/trades

https://nationalpost.com/pmn/sports-pmn/european-talent-powers-overseas-ratings-boom

https://www.si.com/nba/2013/07/10/nba-salary-cap-luxury-tax-figures-2013-14

http://www.ageofbasketball.net/archives/3373

https://www.sportime.gr/basket/nba/nba-draft-i-lotaria-etsi-litourgi/

https://en.wikipedia.org/wiki/NBA_salary_cap

https://en.wikipedia.org/wiki/National_Basketball_Association

ΠΑΡΑΡΤΗΜΑΤΑ

ΠΑΡΑΡΤΗΜΑ Ι: Δεδομένα Πάνελ (Panel Data)

Ι1. Έννοια

Μια ειδική περίπτωση στατιστικών στοιχείων, που αποτελούν συνδυασμό διαστρωματικών στοιχείων και χρονολογικών σειρών (cross – sectional time series) είναι τα panel data. Ονομάζονται, επίσης, και σφαιρικά δεδομένα (longitudinal data) ή διαστρωματικές χρονοσειρές δεδομένων (cross-sectional time series data). Καθώς περιλαμβάνουν επαναλαμβανόμενες παρατηρήσεις, οι οποίες συλλέγονται για μια σειρά χρονικών περιόδων για το ίδιο διαστρωματικό επίπεδο, εκφράζουν μια διαστρωματική μονάδα διαχρονικά.

Ανάλογα με τον αριθμό των χρονικών περιόδων (Τ) και των διαστρωματικών στοιχείων (Ν) διακρίνουμε τις εξής δύο κατηγορίες: (α) short panel, πρόκειται για δεδομένα με μικρό αριθμό χρονικών περιόδων και μεγάλο αριθμό διαστρωματικών στοιχείων και (β) long panel με μικρό αριθμό διαστρωματικών στοιχείων και μεγάλο αριθμό χρονικών περιόδων.

Στη συγκεκριμένη εργασία διαστρωματική μονάδα αποτελεί κάθε ομάδα που αγωνίζεται στο NBA, ενώ η χρονοσειρά αφορά την περίοδο μεταξύ των seasons 2005-2006 και 2019-2020. Τα δεδομένα καταχωρήθηκαν και χρησιμοποιήθηκαν σε short panel data.

Ι2. Το βασικό υπόδειγμα με δεδομένα panel

Το πιο σημαντικό χαρακτηριστικό των μελετών που χρησιμοποιούν δεδομένα panel είναι ότι περιλαμβάνουν δυναμικές αλληλεπιδράσεις στα υποδείγματα. Με άλλα λόγια, παρέχουν εκτιμήσεις για τον τρόπο με τον οποίο μεταβάλλονται τα χαρακτηριστικά των οικονομικών (διαστρωματικών) μονάδων στο πέρασμα του χρόνου, αλλά και μεταξύ αυτών του δείγματος.

Στα δεδομένα panel υπάρχουν δύο μορφές διακύμανσής τους, οι οποίες προσδιορίζουν σε σημαντικό βαθμό τη μέθοδο εκτίμησης του υποδείγματος:

- Η πρώτη μορφή διακύμανσης των στοιχείων πάνελ αποκαλείται διακύμανση εντός (within variation) και σχετίζεται με τη μεταβολή των δεδομένων panel εντός της μεμονωμένης διαστρωματικής μονάδας τη στιγμή που η διακύμανση ανάμεσα στις διαστρωματικές μονάδες παραμένει σταθερή.

- Η δεύτερη μορφή διακύμανσης των δεδομένων πάνελ ονομάζεται διακύμανση μεταξύ (variation within) σχετίζεται με τη μεταβολή της διακύμανσης ανάμεσα στις διαστρωματικές μονάδες. Πιο αναλυτικά, παρατηρείται ότι εντός των μεμονωμένων διαστρωματικών μονάδων η διακύμανση είναι σταθερή, ενώ η διακύμανση ανάμεσα στις διαστρωματικές μονάδες μπορεί να μεταβάλλεται.

Η διανυσματική απεικόνιση των δεδομένων panel παρέχει τη δυνατότητα να εφαρμοστούν διάφορες τεχνικές εκτίμησης των παραμέτρων του υποδείγματος, όπως GLS (Generalized Least Squares), SURE (Seemingly Unrelated Regression), TSLS (Two-Stage Least Squares), GMM (Generalized Method of Moments), FIML (Full Information Maximum Likelihood), 3SLS (Three – Stage Least Squares), κ.λπ. Η τεχνική OLS (Ordinary Least Squares) ακόμα και όταν τα κατάλοιπα είναι ομοσκεδαστικά και δεν συσχετίζονται δεν ενδείκνυται για την εκτίμηση των δεδομένων panel καθώς αγνοεί τις ιδιότητες αυτών με αποτέλεσμα οι συντελεστές να μην είναι ακριβείς.

Το γραμμικό υπόδειγμα panel έχει την εξής μορφή:

$$Y_{it}=a_{it}+\sum_{j=1}^{k} b_j x_{jit}+\varepsilon_{it} \qquad (1)$$

όπου $i = 1, 2, \ldots, N$ το πλήθος των διαστρωματικών μονάδων, $t = 1, 2, \ldots, T$ είναι ο αριθμός των παρατηρήσεων των χρονολογικών σειρών, a_{it} είναι η σταθερά που αντιστοιχεί στη διαστρωματική μονάδα i $(i= 1, 2, \ldots,N)$, b_j είναι ο συντελεστής

συσχέτισης μεταξύ ανεξάρτητης και εξαρτημένης μεταβλητής, ο οποίος δείχνει πόσο μεταβάλλεται η *Y* όταν το *X* αυξάνεται κατά 1 μονάδα, ε_{it} είναι το σφάλμα και, τέλος, $j = 1, 2, \dots k$ είναι ο αριθμός των συντελεστών που αντιστοιχούν στις ανεξάρτητες μεταβλητές X_{jit} .

Όταν ο αριθμός παρατηρήσεων είναι ο ίδιος για κάθε διαστρωματική μονάδα τα δεδομένα panel ονομάζονται ισορροπημένα (balanced panel data), ενώ όταν διαφέρει ανάμεσα στις διαστρωματικές μονάδες τα δεδομένα panel αποκαλούνται μη ισορροπημένα (unbalanced panel data). Στη συγκεκριμένη μελέτη έχουμε ισορροπημένα (balanced) panel data.

I3. Μοντέλο Σταθερών Επιδράσεων (FEM - Fixed Effect Model)

Αν η σταθερά του υποδείγματος (1) διαφέρει από διαστρωματική μονάδα σε διαστρωματική μονάδα, αλλά η τιμή της σταθεράς a_{it} που αναλογεί σε κάθε διαστρωματική μονάδα υποτίθεται ότι παραμένει διαχρονικά αμετάβλητη έχουμε το μοντέλο σταθερών επιδράσεων (FEM - Fixed Effect Model). Με το μοντέλο αυτό (FEM) εξετάζονται οι αιτίες μεταβολής των μεμονωμένων διαστρωματικών μονάδων, με αποτέλεσμα να καθίσταται εφικτή η διάκριση των ιδιαίτερων γνωρισμάτων και χαρακτηριστικών κάθε διαστρωματικής μονάδας.

Το γραμμικό υπόδειγμα FEM παριστάνεται ως εξής:

$$Y_{it}=a_{it}+\sum_{j=1}^{k} b_j x_{jit}+\varepsilon_{it} \qquad (2)$$

Στα πλαίσιο του υποδείγματος FEM, οι εκτιμητές των συντελεστών $\hat{\alpha}_{1t}, \hat{\alpha}_{2t}, \dots, \hat{\alpha}_{it}$ κάθε διαστρωματικής μονάδας διατηρούνται αμετάβλητοι για την εξεταζόμενη χρονική περίοδο.

I4. Μοντέλο Τυχαίων Επιδράσεων (REM - Random Effects Model)

Αν η σταθερά a_{it} είναι τυχαία και όχι αμετάβλητη, το υπόδειγμα ονομάζεται μοντέλο τυχαίων επιδράσεων (Random Effects Model-REM). Ο όρος a_{it} δεν

συσχετίζεται με τις επεξηγηματικές μεταβλητές ενώ ο στοχαστικός όρος ε_{it} παίζει ρυθμιστικό ρόλο κατά τη διαδικασία εκτίμησης του υποδείγματος.

Εφόσον η σταθερά a_{it} είναι τυχαία μεταβλητή σε κάθε διαστρωματική μονάδα θα γράφεται ως: $a_{it} = \mu + u_{it}$ (3) , όπου $i = 1, 2, \ldots, N$. Ο παράγοντας u_{it} είναι ο στοχαστικός όρος και η παράμετρος μ είναι άγνωστη και αντιπροσωπεύει τη μέση τιμή των τυχαίων σταθερών a_{it}. Για το στοχαστικό όρο u_{it} ισχύει $E[u_{it}] = 0$ και Var $(u_{it}) = \sigma_u^2$ ενώ για την παράμετρο μ ισχύει: $E[\mu] = \mu$ και $Var(\mu) = \sigma_\mu^2$.

Εισάγοντας τη σχέση (3) στο βασικό γραμμικό υπόδειγμα λαμβάνουμε το υπόδειγμα REM:

$$Y_{it}=\mu+u_{it}+\sum_{j=1}^{k} b_j x_{jit}+\varepsilon_{it} \quad (4)$$

ή θέτοντας $v_{it}=u_{it}+\varepsilon_{it}$ (5)

προκύπτει το ισοδύναμο υπόδειγμα:

$$Y_{it}=\mu+\sum_{j=1}^{k} b_j x_{jit}+v_{it} \quad (6)$$

όπου πλέον ο στοχαστικός όρος είναι το v_{it} , που αποτελείται από το u_{it} , το οποίο έχει να κάνει με τις μεταβολές των μεμονωμένων διαστρωματικών μονάδων παραμένοντας αμετάβλητος μεταξύ των διαστρωματκών μονάδων με το πέρασμα του χρόνου, και το ε_{it} .

Οι υποθέσεις του υποδείγματος REM είναι οι εξής:

$a_{it} \sim N(0, \sigma_a^2)$, $u_{it} \sim N(0, \sigma_u^2)$, $\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$

$E[u_{it} u_{jt}] = 0$ για $i \neq j$

$E[u_{it} \varepsilon_{it}] = 0$ (7)

$v_{it} \sim N(0, \sigma_v^2)$, $Cov(v_{it} v_{it}) = \sigma_a^2 + \sigma_\varepsilon^2$

$Cov(v_{it} v_{is}) = \sigma_a^2$ για $t \neq s$, $Cov(v_{it} v_{jt})=0$ για $i \neq j$

I5. Υπόδειγμα FEM ή REM

Μεταξύ των υποδειγμάτων σταθερών και τυχαίων επιδράσεων παρουσιάζονται σημαντικές διαφορές. Στο πρώτο υπόδειγμα κάθε διαστρωματική μονάδα έχει τη δική της σταθερά που εκτιμάται με τη μέθοδο των ψευδομεταβλητών. Από την

άλλη, στο δεύτερο υπόδειγμα εκτιμάται η παράμετρος μ είτε με την μέθοδο GLS ή με την ML και εν συνεχεία η σταθερά a_{it} κάθε μεμονωμένης διαστρωματικής μονάδας προσδιορίζεται μέσω του παράγοντα u_{it}.

Για την επιλογή της καταλληλότερης μεθόδου, μεταξύ αυτής των σταθερών ή των τυχαίων επιδράσεων, για την εκτίμηση υποδειγμάτων με δεδομένα panel γίνεται χρήση ελέγχων όπως Hausman (1978).

- **Έλεγχος Hausman**

Αν υποθέσουμε ότι θέλουμε να εκτιμήσουμε το υπόδειγμα $Y_{it} = bX_{it} + \varepsilon_{it}$ (8) , όπου $\varepsilon_{it} = a_{it} + u_{it}$, δηλαδή ο στοχαστικός όρος αποτελείται από τα a_{it} και u_{it} .
Βάση της υπόθεσης ότι ο όρος u_{it} δεν συσχετίζεται με το διάνυσμα των μεταβλητών X_{it} έχουμε $E[X_{it}, u_{it}] = 0$. Ο όρος αυτός ενδέχεται να μεταβάλλεται από παρατήρηση σε παρατήρηση και ανά διαστρωματική μονάδα. Όσον αφορά τον όρο a_{it} ονομάζεται ατομική επίδραση (individual effect) και εκφράζει τη συμπεριφορά κάθε μεμονωμένης διαστρωματικής μονάδας ενώ συντελεί στη διερεύνηση της συμπεριφοράς κάθε διαστρωματικής μονάδας (αν παραμένει σταθερή ή μεταβάλλεται) με το πέρασμα του χρόνου.

Ο έλεγχος Hausman πραγματοποιείται για την απόρριψη ή αποδοχή της μηδενικής υπόθεσης H_0 : a_{it} δεν συσχετίζεται με το διάνυσμα των μεταβλητών X_{it} έναντι της εναλλακτικής H_1 : a_{it} συσχετίζεται με το διάνυσμα των μεταβλητών X_{it} .

Αν από τη διαδικασία του στατιστικού ελέγχου καταλήξουμε σε αποδοχή της μηδενικής υπόθεσης, τότε για την εκτίμηση του υποδείγματος καταλληλότερη είναι η μέθοδος των τυχαίων επιδράσεων. Διαφορετικά, σε περίπτωση απόρριψης της μηδενικής υπόθεσης, καταλληλότερη είναι η προσέγγιση των σταθερών επιδράσεων για το υπόδειγμά μας.

I6. Δυναμικά Υποδείγματα με Δεδομένα Panel (Dynamic Panel Data Model – DPDM)

Στα δυναμικά υποδείγματα με δεδομένα panel (DPDM) στην ομάδα των ερμηνευτικών μεταβλητών συμπεριλαμβάνονται μεταβλητές με χρονικές υστερήσεις της εξαρτημένης. Στα υποδείγματα DPDM οι μεταβλητές εκφράζονται σε επίπεδα ή σε πρώτες διαφορές. Οι ακόλουθες διατυπώσεις απεικονίζουν τις δύο διαφορετικές μορφές υποδειγμάτων αντιστοίχως:

$$Y_{it} = a_{it} + \sum_{j=1}^{p} \gamma_j Y_{it-1} + X_{it} b + \varepsilon_{it} \quad (9)$$

και

$$\Delta Y_{it} = a_{it} + \sum_{j=1}^{p} \gamma_j Y_{it-1} + \Delta X_{it} b + \Delta\varepsilon_{it} \quad (10)$$

Παρακάτω θα δούμε ότι στην παρούσα μελέτη θα εισαχθεί στο μοντέλο χρονική υστέρηση της εξαρτημένης μεταβλητής σύμφωνα με την πρώτη διατύπωση.

I7. Έλεγχος μοναδιαίας ρίζας

Καταρχήν, μία χρονοσειρά χαρακτηρίζεται ως στάσιμη (stationary) αν τα στατιστικά χαρακτηριστικά παραμένουν αμετάβλητα διαχρονικά. Με άλλα λόγια, η στασιμότητα σημαίνει ότι οι μέσοι και οι διακυμάνσεις της χρονικής σειράς είναι ίδια για όλες τις τιμές της και η συνδιακύμανση μεταξύ δύο χρονικών περιόδων, t και t+k, εξαρτάται μόνο από την υστέρηση μεταξύ των δύο αυτών χρονικών περιόδων.

Για τον έλεγχο στασιμότητας χρησιμοποιείται ο έλεγχος μοναδιαίας ρίζας (the unit root test) ο οποίος εξετάζει αν η προς μελέτη μεταβλητή έχει τη τάση να επιστρέφει στο μέσο (mean reversion). Η αφετηρία των ελέγχων αυτών αποτελεί το εξής αυτοπαλίνδρομο υπόδειγμα πρώτης τάξης:

$$y_{i,t} = \rho_i y_{i,t-1} + z'_{i,t}\gamma_i + \varepsilon_{it} \quad (11)$$

όπου το ε_{it} είναι ένα στάσιμος όρος σφάλματος, το z'_{it} λαμβάνοντας την τιμή: (α)0 δεν περιέχει ούτε ατομικές (χωρικές) ούτε χρονικές επιδράσεις, (β)1 περιέχει μόνο ατομικές επιδράσεις και (γ) (1,t) περιέχει ατομικές επιδράσεις και μία σταθερή χρονική τάση.

Η παραπάνω σχέση γράφεται ισοδύναμα

$$\Delta y_{i,t} = \varphi_i y_{i,t-1} + z'_{i,t}\gamma_i + \varepsilon_{it} \quad (12)$$

όπου $\phi_i = \rho_i - 1$. Στην περίπτωση: (α) $|\rho_i| < 1$ το Y χαρακτηρίζεται ως ελαφρώς στάσιμη χρονοσειρά, (β) $\rho_i = 1$ το Y περιέχει μια μοναδιαία ρίζα.

Όσον αφορά τα data panel, έχουν αναπτυχθεί ποικίλες τεχνικές ελέγχου για τα διαστρωματικά και διαχρονικά στοιχεία οι οποίοι διακρίνονται σε πρώτης και δεύτερης γενιάς. Οι έλεγχοι πρώτης γενιάς, όπως των Harris και Tzavalis (1999), Levin, Lin και Chu (2002), των IM, Pesaran και Shin (2003), προϋποθέτουν την ανεξαρτησία των μονάδων των διαστρωματικών στοιχείων. Από την άλλη πλευρά, εκείνοι της δεύτερης γενιάς, όπως του Fisher, των Chang (2002,2004) Bai και Ng(2004), Phillips και Sul(2003), Moon and Perron (2004a), Choi (2002) και Pesaran (2003) μεταξύ άλλων, απορρίπτουν την υπόθεση ανεξαρτησίας.

Οι περισσότεροι έλεγχοι μοναδιαίας ρίζας σε δεδομένα panel χρησιμοποιούν την ακόλουθη γενική εξίσωση:

$$\Delta y_{i,t} = \rho_i y_{i,t-1} + \sum_{L=1}^{p} \theta_{i,L} \Delta y_{i,t-L} + a_{ki} d_{kt} + \varepsilon_{it} \quad (13)$$

όπου p το μήκος της χρονικής υστέρησης, d_{kt} είναι ορισμένες μεταβλητές με d_{1t} = {null}, d_{2t} = {1}, d_{3t} = {1,t} και α_{kt} οι εκτιμητές των συντελεστών τους. Η μηδενική υπόθεση H_0: $\rho_i = 0$ εκφράζει την ύπαρξη μοναδιαίας ρίζας για κάθε μεμονωμένη διαστρωματική μονάδα i, ενώ η εναλλακτική υπόθεση H_1: $\rho_i < 0$ υποδεικνύει μια στάσιμη διαδικασία.

Οι Levin, Lin και Chu (2002) υποστήριξαν ότι δεδομένου ότι το ρ_i είναι σταθερό για κάθε i, η εναλλακτική υπόθεση είναι ότι τα ρ_i είναι πανομοιότυπα και αρνητικά. Οι Harris και Tzavalis (1999) πρότειναν έναν παρόμοιο αλλά πιο «ευέλικτο» έλεγχο όταν η χρονική περίοδος, T, είναι σχετικά μικρή, με μηδενική υπόθεση ύπαρξης μιας μοναδιαίας ρίζας και εναλλακτική την ύπαρξη μοναδικής στάσιμης τιμής.

ΠΑΡΑΡΤΗΜΑ ΙΙ: Επαυξημένος Έλεγχος Μοναδιαίας Ριζας των Dickey-Fuller (ADF Test)

Θεωρούμε το μοντέλο: $x_t = \theta x_{t-1} + \varepsilon_t$ (1)

Τα κύρια χαρακτηριστικά του Dickey – Fuller (1979) ελέγχου είναι ότι περιλαμβάνει μια χρονική υστέρηση με σκοπό να ερευνηθεί πότε υπάρχει μοναδιαία ρίζα ή όχι. Αφαιρώντας το x_{t-1} και από τα δύο μέρη της παραπάνω ισότητας καταλήγουμε στην εξής σχέση:

$$x_t - x_{t-1} = (\theta-1)x_{t-1} + \varepsilon_t \;\; => \Delta x_t = \delta x_{t-1} + \varepsilon_t \qquad (2)$$

Η μηδενική υπόθεση του ελέγχου είναι $Ho : \delta=0$ και η εναλλακτική $H1: \delta\neq 0$.

Εάν $\delta=0$ τότε $\theta=1$, έχουμε την ύπαρξη μοναδιαίας ρίζας και συνεπώς αντιμετωπίζουμε το πρόβλημα της μη στασιμότητας στις μεταβλητές του μοντέλου μας. Θα πρέπει να αναφερθεί ότι η τιμή t-statistic του εκτιμημένου συντελεστή δεν ακολουθεί την συμβατική student-t κατανομή. Γι' αυτό τον λόγο χρησιμοποιούνται συγκεκριμένες κριτικές τιμές από τον Dickey –Fuller για την απόρριψη ή όχι της μηδενικής υπόθεσης. Μια από τις βασικές υποθέσεις του Dickey –Fuller τεστ είναι ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ των τυπικών σφαλμάτων. Στην πράξη όμως συνήθως παρατηρείται αυτοσυσχέτιση μεταξύ των τυπικών σφαλμάτων κάτι που οδηγεί στην δημιουργία ενός πιο ανεπτυγμένου τεστ γνωστό ως Augmented Dickey –Fuller test (ADF) το οποίο υπερβαίνει το πρόβλημα αυτό. Σύμφωνα με το ADF τεστ η σχέση της παλινδρόμησης μπορεί να μετατραπεί και να γραφτεί στις τρεις ακόλουθες μορφές:

$$\Delta x_t = \delta x_{t-1} + \alpha_1 \Delta x_{t-1} + \ldots + \alpha_q \Delta x_{t-q} + \varepsilon_t \qquad (3)$$

$$\Delta x_t = \beta_0 + \delta x_{t-1} + \alpha_1 \Delta x_{t-1} + \ldots + \alpha_q \Delta x_{t-q} + \varepsilon_t \qquad (4)$$

$$\Delta x_t = \beta_0 + \beta_1 t + \delta x_{t-1} + \alpha_1 \Delta x_{t-1} + \ldots + \alpha_q \Delta x_{t-q} + \varepsilon_t \qquad (5)$$

Στην παραπάνω σχέση το x_t και οι χρονικές υστερήσεις αντιπροσωπεύουν τις μεταβλητές Y,P κλπ. Ο αριθμός των παρατηρήσεων στο δείγμα παίζει τον ρόλο της τάσης και ο όρος β_o αντιπροσωπεύει την σταθερά . Το ε_t θεωρείται ότι είναι το τυχαίο σφάλμα. Μπορούμε να παρατηρήσουμε ότι ούτε η σταθερά , ούτε η τάση

περιλαμβάνεται στην σχέση (1). Η (3) σχέση περιλαμβάνει σταθερά αλλά όχι τάση , ενώ η (4) σχέση περιλαμβάνει και τάση και σταθερά. Ο λόγος που χρησιμοποιούνται οι διαφορές στις χρονικές υστερήσεις q, είναι για να βεβαιωθούμε ότι δεν υπάρχει καμία γραμμική συσχέτιση μεταξύ των εκτιμημένων όρων των τυπικών σφαλμάτων . Η σταθερά και η τάση χρειάζεται να ελεγχθούν πότε παρέχουν πρόσθετη ισχύ. Με αυτόν τον τρόπο οι μη σημαντικές μεταβλητές (σταθερά, τάση) αφαιρούνται από το υπόδειγμα. Άρα υπάρχουν τρία πιθανά μοντέλα που μπορούν αν ελεγχθούν για τη σχέση (5):

1. Αν $\beta_0=0, \beta_1=0$ και $\delta=0$ η (4) συνεπάγεται το υπόδειγμα $\Delta x_t = \alpha_1 \Delta x_{t-1} + \varepsilon_t$
2. Αν $\beta_1=0$ και $\delta=0$ η (4) συνεπάγεται το υπόδειγμα $\Delta x_t = \beta_0 + \alpha_1 \Delta x_{t-1} + \varepsilon_t$
3. Αν $\beta_1 \neq 0$ η (4) συνεπάγεται το υπόδειγμα $\Delta x_t = \beta_0 + \beta_1 t + \alpha_1 \Delta x_{t-1} + \varepsilon_t$

Εάν η t- value για την μεταβλητή που έχουμε επιλέξει είναι μεγαλύτερη από την ADF κριτική τιμή , η μηδενική υπόθεση (δ=0) απορρίπτεται και η μεταβλητή θεωρείται στάσιμη. Διαφορετικά σε περίπτωση που η μηδενική υπόθεση δεν απορριφθεί και η μεταβλητή δεν είναι στάσιμη , θα πρέπει να ελεγχθούν με τη σειρά οι πρώτες και δεύτερες διαφορές , μέχρι να επιτευχθεί η στασιμότητα. Αυτή η διαδικασία πρέπει να επαναληφθεί για όλες τις ανεξάρτητες μεταβλητές του μοντέλου.

ΠΑΡΑΡΤΗΜΑ ΙΙΙ: Προηγμένα Στατιστικά (Advanced Stats)

Οι προηγμένες στατιστικές στο μπάσκετ αναφέρονται στην ανάλυση στατιστικών μπάσκετ μέσω αντικειμενικών «αποδεικτικών» στοιχείων.

Ένα βασικό δόγμα για πολλούς σύγχρονους αναλυτές μπάσκετ είναι ότι το μπάσκετ αξιολογείται καλύτερα στο επίπεδο των κατοχών (possessions). Κατά τη διάρκεια ενός μόνο παιχνιδιού, και οι δύο ομάδες έχουν περίπου τον ίδιο αριθμό κατοχών, επειδή εναλλάσσουν την κατοχή. Ωστόσο, κατά τη διάρκεια της σεζόν, οι ομάδες παίζουν με πολύ διαφορετικούς ρυθμούς, οι οποίοι μπορούν να επηρεάσουν δραματικά τους πόντους που σημείωσαν και τους πόντους που δέχονται ανά παιχνίδι. Ως εκ τούτου, αυτοί οι αναλυτές ευνοούν τη χρήση των πόντων που βαθμολογούνται ανά 100 κατοχές (επιθετική βαθμολογία) και των πόντων που επιτρέπονται ανά 100 κατοχές (αμυντική βαθμολογία).

Ένα δεύτερο βασικό δόγμα είναι ότι τα στατιστικά στοιχεία ανά λεπτό είναι πιο χρήσιμα για την αξιολόγηση των παικτών από τα στατιστικά στοιχεία ανά παιχνίδι.

Τα προηγμένα (advanced) και παραδοσιακά (traditional) στατιστικά που χρησιμοποιήθηκαν στην παρούσα έρευνα είναι τα εξής:

1. eFG% (Effective Field Goal Percentage) εκφράζει το ποσοστό ευστοχίας για λογαριασμό του γεγονότος πως η προσπάθεια για τρίποντο μετράει για τρεις πόντους, ενώ η κανονική προσπάθεια μετράει μόνο για δύο πόντους. Ο τύπος υπολογισμού είναι: eFG% = ((FGM + (0.5 * 3PM)) / FGA, όπου FGM είναι τα εύστοχα σουτ, 3PM τα εύστοχα τρίποντα και FGA το πλήθος των προσπαθειών για σουτ άστοχων και εύστοχων
2. Assist Ratio είναι το ποσοστό των κατοχών μιας ομάδας που καταλήγει σε μια assist (δηλαδή πάσα που οδηγεί αμέσως σε εύστοχο καλάθι). Ο τύπος υπολογισμού είναι: Assist Ratio = (Assists)*100 / [(Field Goal Attempts) + (Free Throw Attempts* 0.44) +(Assists)+(Turnovers)]
3. OREB% (Offensive Rebounds Percentage) είναι το ποσοστό των επιθετικών rebound που αποκτά ένας παίκτης ή μία ομάδα από τα διεκδικήσιμα ή διαθέσιμα επιθετικά rebound

4. DREB% (Defensive Rebounds Percentage) είναι το ποσοστό των αμυντικών rebound που αποκτά ένας παίκτης ή μία ομάδα από τα διεκδικήσιμα ή διαθέσιμα αμυντικά rebound
5. TOV% (Turnovers Percentage) είναι το ποσοστό των κατοχών μιας ομάδας που καταλήγουν σε λάθος
6. PACE είναι το πλήθος των κατοχών μιας ομάδας ανά 48 λεπτά που διαρκεί ένας αγώνας. Εκφράζει τον ρυθμό του παιχνιδιού μιας ομάδας, καθώς όσο πιο πολλές οι κατοχές μιας ομάδας ανά 48 λεπτά τόσο πιο γρήγορες επιθέσεις πραγματοποιεί η ομάδα και αναλόγως όσο λιγότερες οι κατοχές τόσο πιο αργές οι επιθέσεις.

ΠΑΡΑΡΤΗΜΑ IV: Οι ομάδες του NBA

Πίνακας Π1: Ομάδες NBA

Team	id	Full Name of Team
BOS	1	Boston Celtics
BKN	2	Brooklyn Nets
NYK	3	New York Knicks
PHI	4	Philadelphia 76ers
TOR	5	Toronto Raptors
CHI	6	Chicago Bulls
CLE	7	Cleveland Cavaliers
DET	8	Detroit Pistons
IND	9	Indiana Pacers
MIL	10	Milwaukee Bucks
ATL	11	Atlanta Hawks
CHA	12	Charlotte Hornets
MIA	13	Miami Heat
ORL	14	Orlando Magic
WAS	15	Washington Wizards
DEN	16	Denver Nuggets
MIN	17	Minnesota Timberwolves
OKC	18	Oklahoma City Thunder
POR	19	Portland Trail Blazers
UTA	20	Utah Jazz
GSW	21	Golden State Warriors
LAC	22	Los Angeles Clippers
LAL	23	Los Angeles Lakers
PHX	24	Phoenix Suns
SAC	25	Sacramento Kings
DAL	26	Dallas Mavericks
HOU	27	Houston Rockets
MEM	28	Memphis Grizzlies
NOP	29	New Orleans Pelicans
SAS	30	San Antonio Spurs

ΠΑΡΑΡΤΗΜΑ V: Μήτρα Βαρών (W) και Κανονικοποιημένη Μήτρα Βαρών (Wnorm)

Πίνακας Π2: Μήτρα Βαρών

W	BOS	BKN	NYK	PHI	TOR	CHI	CLE	DET	IND	MIL	ATL	CHA	MIA	ORL	WAS	DEN	MIN	OKC	POR	UTA	GSW	LAC	LAL	PHX	SAC	DAL	HOU
BOS	0,0	1,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
BKN	1,0	0,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
NYK	1,0	1,0	0,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
PHI	1,0	1,0	1,0	0,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
TOR	1,0	1,0	1,0	1,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CHI	0,5	0,5	0,5	0,5	0,5	0,0	1,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CLE	0,5	0,5	0,5	0,5	0,5	1,0	0,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
DET	0,5	0,5	0,5	0,5	0,5	1,0	1,0	0,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
IND	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	0,0	1,0	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
MIL	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	1,0	0,0	0,5	0,5	0,5	0,5	0,5	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ATL	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CHA	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	0,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
MIA	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0	0,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ORL	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	0,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
WAS	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	1,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
DEN	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5
MIN	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	0,0	1,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5
OKC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	0,0	1,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5
POR	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	0,0	1,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5
UTA	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	1,0	1,0	1,0	1,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5
GSW	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,0	1,0	1,0	1,0	1,0	0,5	0,5
LAC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	1,0	0,0	1,0	1,0	1,0	0,5	0,5
LAL	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	1,0	1,0	0,0	1,0	1,0	0,5	0,5
PHX	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	0,0	1,0	0,5	0,5
SAC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	1,0	1,0	1,0	1,0	0,0	0,5	0,5
DAL	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,0	1,0
HOU	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	0,0
MEM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0
NOP	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0
SAS	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	0,5	1,0	1,0

Πίνακας Π3: Κανονικοποιημένη Μήτρα Βαρών

Wnorm	BOS	BKN	NYK	PHI	TOR	CHI	CLE	DET	IND	MIL	ATL	CHA	MIA	ORL	WAS	DEN	MIN	OKC	POR	UTA	GSW	LAC	LAL	PHX	SAC	DAL
BOS	0,0	0,11	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
BKN	0,11	0,0	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
NYK	0,11	0,11	0,0	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
PHI	0,11	0,11	0,11	0,0	0,11	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
TOR	0,11	0,11	0,11	0,11	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CHI	0,06	0,06	0,06	0,06	0,06	0,0	0,11	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CLE	0,06	0,06	0,06	0,06	0,06	0,11	0,0	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
DET	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,0	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
IND	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,0	0,11	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
MIL	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,11	0,0	0,06	0,06	0,06	0,06	0,06	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ATL	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0	0,11	0,11	0,11	0,11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
CHA	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,0	0,11	0,11	0,11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
MIA	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,0	0,11	0,11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
ORL	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,0	0,11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
WAS	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,11	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0
DEN	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,11	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06
MIN	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,11	0,0	0,11	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06
OKC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,11	0,11	0,0	0,11	0,11	0,06	0,06	0,06	0,06	0,06	0,06
POR	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,11	0,11	0,11	0,0	0,11	0,06	0,06	0,06	0,06	0,06	0,06
UTA	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,11	0,11	0,11	0,11	0,0	0,06	0,06	0,06	0,06	0,06	0,06
GSW	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,0	0,11	0,11	0,11	0,11	0,06
LAC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,11	0,0	0,11	0,11	0,11	0,06
LAL	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,0	0,11	0,11	0,06
PHX	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,0	0,11	0,06
SAC	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,11	0,11	0,11	0,11	0,0	0,06
DAL	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,0
HOU	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11
MEM	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11
NOP	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11
SAS	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,0	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,06	0,11