

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ



ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΚΑΙ ΑΣΦΑΛΙΣΤΙΚΗΣ ΕΠΙΣΤΗΜΗΣ

ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ ΣΠΟΥΔΩΝ

ΣΤΗΝ ΕΦΑΡΜΟΣΜΕΝΗ ΣΤΑΤΙΣΤΙΚΗ

**ΟΙΚΟΝΟΜΙΚΕΣ ΑΝΙΣΟΤΗΤΕΣ ΜΕΤΑΞΥ ΤΩΝ
ΗΛΙΚΙΩΜΕΝΩΝ ΣΕ ΧΩΡΕΣ ΤΗΣ ΕΥΡΩΠΑΪΚΗΣ ΕΝΩΣΗΣ:
ΑΝΑΛΥΣΗ ΒΑΣΕΙ ΤΩΝ ΜΙΚΡΟΔΕΔΟΜΕΝΩΝ ΤΟΥ
ΠΡΟΓΡΑΜΜΑΤΟΣ SHARE**

Μαρία Μαραντίδου

Διπλωματική εργασία

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

ΠΕΙΡΑΙΑΣ

ΟΚΤΩΒΡΙΟΣ 2017

UNIVERSITY OF PIRAEUS



DEPARTMENT OF STATISTIC AND INSURANCE SCIENCE

POSTGRADUATE PROGRAMM IN APPLIED STATISTICS

**ECONOMIC INEQUALITIES AMONG
ELDERLY IN EUROPEAN COUNTRIES:
AN ANALYSIS BASED ON SHARE MICRODATA**

by

Maria Marantidou

MSc Dissertation

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

PIRAEUS, GREECE

OCTOBER 2017

ΕΥΧΑΡΙΣΤΙΕΣ

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή μου κ. Κλέωνα Τσίμπο για την πολύτιμη βοήθειά του και την καθοδήγησή του κατά τη διάρκεια εκπόνησης της διπλωματικής μου εργασίας καθώς επίσης και την Αναπληρώτρια καθηγήτρια κ. Γεωργία Βερροπούλου και τον Επίκουρο καθηγητή κ. Χαράλαμπο Ευαγγελάρα. Όμως περισσότερο από όλους θα ήθελα να ευχαριστήσω την οικογένεια μου για την υποστήριξη της, την υπομονή της και την πίστη της σε εμένα.

ΠΕΡΙΛΗΨΗ

Τα τελευταία χρόνια ολοένα και περισσότερες έρευνες εστιάζονται στη μελέτη της ύπαρξης αλληλεπίδρασης μεταξύ των τομέων υγείας, κοινωνικοοικονομικής κατάστασης και συνταξιοδότησης. Η παρούσα διπλωματική εργασία στοχεύει στην εφαρμογή καθιερωμένων μεθόδων με σκοπό τη διερεύνηση κοινωνικοοικονομικών ανισοτήτων μεταξύ ατόμων ηλικίας 50 ετών και άνω σε χώρες της Ευρωπαϊκής Ένωσης. Η ανάλυση βασίζεται στο δεύτερο κύμα συλλογής στοιχείων του προγράμματος SHARE (Survey of Health, Ageing and Retirement in Europe) και αφορά 14 χώρες. Για τη στατιστική ανάλυση χρησιμοποιούνται οικονομικές μεταβλητές, μεταβλητές υγείας καθώς και δημογραφικές μεταβλητές.

Αρχικά αναφέρονται οι μέθοδοι εκτίμησης των κοινωνικοοικονομικών ανισοτήτων εστιάζοντας στην Καμπύλη Lorenz, στο Δείκτη Συγκέντρωσης και στο εύρος των πιθανών τιμών του καθώς και η Λογιστική Παλινδρόμηση. Στη συνέχεια, εξάγονται βασικά περιγραφικά μέτρα του δείγματος και εξετάζεται η ύπαρξη επίδρασης μεταξύ των μεταβλητών υγείας σε σχέση με τις οικονομικές μεταβλητές μέσω του Δείκτη Συγκέντρωσης καθώς επίσης και οι επιδράσεις δημογραφικών και οικονομικών παραγόντων στον τομέα της υγείας μέσω της Λογιστικής Παλινδρόμησης. Όσον αφορά την τελευταία μέθοδο εστιάζεται, πέραν από το σύνολο των χωρών και στην Ελλάδα. Τέλος, εξάγονται κάποια συμπεράσματα για τις παραπάνω επιδράσεις, μελετώντας το συνδυασμό των μεθόδων ανάλυσης που χρησιμοποιούνται στην εργασία.

Λέξεις - Κλειδιά: Κοινωνικοοικονομικές ανισότητες, Δείκτης Συγκέντρωσης, Λογιστική Παλινδρόμηση

ABSTRACT

In recent years, more and more research has focused on studying the existence of associations between health, socio-economic status and retirement. This dissertation aims at applying established methods to investigate socio-economic inequalities among people aged 50 and over in European countries. The analysis is based on the second wave of SHARE (Survey of Health, Ageing and Retirement in Europe) data collection and concerns 14 countries. Economic, health and demographic variables are used for statistical analysis.

Initially the methods of assessing socio-economic inequalities are mentioned, focusing on the Lorenz Curve, the Concentration Index and the range of possible values, as well as Logistic Regression. Subsequently, basic descriptive statistics of the sample are presented and the effect between health and economic variables using the concentration index is examined, as well as the effects of demographic and economic factors on health indicators using Logistic Regression. With regard to this latter method, apart from all the countries it also focuses on Greece too. Finally, some conclusions are derived from the above effects, taking into account the combination of the analysis methods used in this paper.

Keywords: Socio-economics inequalities, Concentration Index, Logistic Regression

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΥΧΑΡΙΣΤΙΕΣ	iii
ΠΕΡΙΛΗΨΗ	iv
ABSTRACT.....	v
ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ	viii
ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ.....	x
1. ΕΙΣΑΓΩΓΗ	1
2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ	3
2.1 ΤΟ ΕΡΕΥΝΗΤΙΚΟ ΠΡΟΓΡΑΜΜΑ SHARE.....	3
2.2 ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΗΣ ΚΟΙΝΩΝΙΚΟΟΙΚΟΝΟΜΙΚΩΝ ΑΝΙΣΟΤΗΤΩΝ	6
2.2.1 ΚΑΜΠΥΛΗ LORENZ.....	6
2.2.2 ΔΕΙΚΤΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ.....	7
2.2.3 ΕΥΡΟΣ ΠΙΘΑΝΩΝ ΤΙΜΩΝ ΔΕΙΚΤΗ ΣΥΓΚΕΝΤΡΩΣΗΣ.....	8
2.3 ΜΕΘΟΔΟΛΟΓΙΚΟ ΠΛΑΙΣΙΟ.....	9
2.3.1 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ	9
2.3.2 ΠΡΟΣΑΡΜΟΓΗ ΜΟΝΤΕΛΟΥ	10
3. ΔΕΔΟΜΕΝΑ ΕΦΑΡΜΟΓΩΝ	12
3.1 ΜΕΤΑΒΛΗΤΕΣ.....	12
3.2 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ	15
4. ΚΟΙΝΩΝΙΚΟΟΙΚΟΝΟΜΙΚΕΣ ΑΝΙΣΟΤΗΤΕΣ ΚΑΙ ΥΓΕΙΑ.....	22
4.1 ΔΕΙΚΤΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ.....	22
4.2 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ	34
5. ΣΥΜΠΕΡΑΣΜΑΤΑ	52
ΠΑΡΑΡΤΗΜΑ.....	55
ΒΙΒΛΙΟΓΡΑΦΙΑ	75

ΚΑΤΑΛΟΓΟΣ ΠΙΝΑΚΩΝ

Πίνακας 3.1 Πίνακας Μεταβλητών	12
Πίνακας 3.2 Πίνακας συχνότητας μεταβλητής country	16
Πίνακας 3.3 Πίνακας συχνότητας μεταβλητής chronic2w2	17
Πίνακας 3.4 Πίνακας συχνότητας μεταβλητής eurodcat.....	18
Πίνακας 3.5 Πίνακας συχνότητας μεταβλητής gali	18
Πίνακας 3.6 Πίνακας συχνότητας για τη μεταβλητή sphus	19
Πίνακας 3.7 Περιγραφικά μέτρα των οικονομικών μεταβλητών του πληθυσμού.....	20
Πίνακας 4.1 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhinc	22
Πίνακας 4.2 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhinc.....	24
Πίνακας 4.3 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhinc	25
Πίνακας 4.4 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή gali όσον αφορά τη μεταβλητή hhinc	27
Πίνακας 4.5 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhnetw	28
Πίνακας 4.6 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhnetw	30
Πίνακας 4.7 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhnetw	31
Πίνακας 4.8 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή gali όσον αφορά τη μεταβλητή hhnetw	33
Πίνακας 6.1 Κωδικοποίηση του “CC” για τη μεταβλητή mergeid.....	55
Πίνακας 6.2 Κωδικοποίηση χωρών για τη μεταβλητή country	55
Πίνακας 6.3 Κωδικοποίηση μεταβλητής chronic2w2.....	55
Πίνακας 6.4 Στοιχεία της μεταβλητής υγείας eurodcat	55
Πίνακας 6.5 Κωδικοποίηση μεταβλητής gender	56
Πίνακας 6.6 Κωδικοποίηση μεταβλητής ageclass	56
Πίνακας 6.7 Κωδικοποίηση της μεταβλητής sphus	56
Πίνακας 6.8 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους Έλληνες ερωτώμενους και τη μεταβλητή hhnetw	56
Πίνακας 6.9 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc	57
Πίνακας 6.10 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους Έλληνες ερωτώμενους και τη μεταβλητή hhnetw	58
Πίνακας 6.11 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc	59
Πίνακας 6.12 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας sphus για τους Έλληνες ερωτώμενους και τη μεταβλητή hhnetw	59
Πίνακας 6.13 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας sphus για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc	60

Πίνακας 6.14 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους Έλληνες ερωτώμενους και τη μεταβλητή hhnetw	61
Πίνακας 6.15 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc	61
Πίνακας 6.16 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw	63
Πίνακας 6.17 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw	64
Πίνακας 6.18 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας sphus για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw	66
Πίνακας 6.19 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw	67
Πίνακας 6.20 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhinc	69
Πίνακας 6.21 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhinc	70
Πίνακας 6.22 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας sphus για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhinc	72
Πίνακας 6.23 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhinc	73

ΚΑΤΑΛΟΓΟΣ ΔΙΑΓΡΑΜΜΑΤΩΝ

Διάγραμμα 3.1 Κατανομή συχνοτήτων ηλικιακών ομάδων	16
Διάγραμμα 3.2 Πυραμίδα ηλικιών.....	17
Διάγραμμα 3.3 Πυραμίδα χρόνων εκπαίδευσης	19
Διάγραμμα 3.4 Θηκόγραμμα μεταβλητής hhnetw συγκριτικά με τη μεταβλητή country	21
Διάγραμμα 3.5 Θηκόγραμμα μεταβλητής hhinc συγκριτικά με τη μεταβλητή country	21
Διάγραμμα 4.1 Δείκτης Συγκέντρωσης για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhinc	23
Διάγραμμα 4.2 Δείκτης Συγκέντρωσης για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhinc.....	24
Διάγραμμα 4.3 Δείκτης Συγκέντρωσης για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhinc	26
Διάγραμμα 4.4 Δείκτης Συγκέντρωσης για τη μεταβλητή gali όσον αφορά τη μεταβλητή hhinc	27
Διάγραμμα 4.5 Δείκτης Συγκέντρωσης για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhnetw....	29
Διάγραμμα 4.6 Δείκτης Συγκέντρωσης για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhnetw..	30
Διάγραμμα 4.7 Δείκτης Συγκέντρωσης για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhnetw	32
Διάγραμμα 4.8 Δείκτης Συγκέντρωσης για τη μεταβλητή gali όσον αφορά τη μεταβλητή hhnetw.....	33

1. ΕΙΣΑΓΩΓΗ

Τα τελευταία χρόνια δίνεται μεγάλη έμφαση στη μελέτη των τομέων υγείας, κοινωνικοοικονομικής κατάστασης και συνταξιοδότησης καθώς και στις αλληλεπιδράσεις μεταξύ αυτών. Ολοένα και περισσότερες έρευνες, κυρίως Ευρωπαϊκού χαρακτήρα, πραγματοποιούνται και συλλέγουν δεδομένα για την ανάλυση της ποιότητας ζωής και των κοινωνικοοικονομικών χαρακτηριστικών των ατόμων ηλικίας 50 ετών και άνω. Μια τέτοιου είδους έρευνα είναι το SHARE, η οποία είναι πανευρωπαϊκού επιπέδου και έχει ως στόχο τη συλλογή και ανάλυση δεδομένων που σχετίζονται με την υγεία, τη γήρανση και τη συνταξιοδότηση με σκοπό τη μελέτη οικονομικών και κοινωνικών παραγόντων καθώς και της υγείας μεταξύ των ατόμων ηλικίας άνω των 50 ετών.

Το SHARE επιτρέπει συγκρίσεις μεταξύ των κρατών-μελών της Ευρωπαϊκής Ένωσης μέσω της διεξαγωγής ερευνών αναφορικά με τις καταστάσεις, τις συμπεριφορές και τις προτιμήσεις που παρατηρούνται στο δεύτερο ήμισυ της ζωής καθώς και τις επιπτώσεις της πληθυσμιακής γήρανσης. Το πρώτο κύμα του προγράμματος SHARE ξεκίνησε το 2004 με 27.000 άτομα από 11 χώρες. Ακολούθησαν άλλα πέντε, με το τελευταίο κύμα (το έκτο) να ολοκληρώνεται το 2015 περιλαμβάνοντας 18 χώρες.

Στις σημερινές κοινωνίες, η γήρανση του πληθυσμού αποτελεί σημαντικό δημογραφικό πρόβλημα. Καθώς αυξάνεται η ηλικία παρατηρείται φθορά του οργανισμού όσον αφορά την κατάσταση της υγείας του και αύξηση των δυσκολιών που αντιμετωπίζουν τα άτομα σε δραστηριότητες. Σε συνδυασμό με την κοινωνική και οικονομική κατάσταση του πληθυσμού αποτελούν πηγή έρευνας για τις συνθήκες διαβίωσης καθώς και για τα μέτρα που πρέπει να λάβει κάθε χώρα προς αποφυγήν υψηλής νοσηρότητας και θνησιμότητας του πληθυσμού της.

Επομένως, ένα σημαντικό ζήτημα που αφορά τη γήρανση συνδέεται με την οικονομική κατάσταση (συγκεκριμένα τη φτώχεια), με το κατά πόσο αυτή επηρεάζει την υγεία του γηραιού πληθυσμού και πώς μεταβάλλεται στις διάφορες ηλικιακές φάσεις της ζωής του ανθρώπου. Όσο αυξάνεται η ηλικία, η υποκειμενική αίσθηση της φτώχειας εξαρτάται ολοένα και περισσότερο από παράγοντες πέραν του εισοδήματος. Γι αυτό το λόγο είναι απαραίτητο να ξεκαθαρίσει η έννοια της φτώχειας καθώς και να προσδιοριστούν οι παράγοντες που επηρεάζουν την κατάσταση της υγείας του γηραιού πληθυσμού.

Η κατάσταση της υγείας ενός ατόμου κρίνεται κυρίως από δύο διαστάσεις: την αντικειμενική, δηλαδή ο γιατρός κρίνει αν ένα άτομο είναι υγιές ή ασθενές μέσω κάποιων εξετάσεων και τη σοβαρότητα της νόσου που μπορεί να έχει, και την υποκειμενική, δηλαδή την προσωπική εκτίμηση της νοσηρότητας των ατόμων κυρίως μέσω κάποιων ψυχολογικών παραγόντων (κατάθλιψη, μοναξιά, κοινωνική απομάκρυνση κ.α.) καθώς και κινητικών και νοητικών ανικανοτήτων.

Από την άλλη μεριά, η φτώχεια σαν έννοια παίζει σημαντικό ρόλο στο σχηματισμό της κοινωνικής πολιτικής. Στις αναπτυγμένες χώρες, χαρακτηρίζεται συνήθως ως η κατάσταση στην οποία ένα άτομο δεν είναι διαθέσιμο να συμμετέχει πλήρως σε κοινωνικά αποδεκτές δραστηριότητες όπως η ζωή στο κοινωνικό σύνολο. Ορίζεται δηλαδή, ως η οικονομική κατάσταση που χαρακτηρίζεται από έλλειψη επαρκών πόρων για την ικανοποίηση βασικών αναγκών. Οι βασικές ανάγκες κοινωνικής ευημερίας ενός ατόμου δεν είναι μόνο τα υλικά αγαθά της αγοράς αλλά και η καλή κατάσταση υγείας, η πρόσβαση στις κοινωνικές υπηρεσίες, στις υπηρεσίες φροντίδας, στις αστικές συγκοινωνίες, στα κοινωνικά δίκτυα καθώς και σε άλλες μη εμφανείς δραστηριότητες.

Ένα από τα σημαντικότερα ζητήματα της παγκόσμιας πολιτικής είναι η μείωση των κοινωνικοοικονομικών ανισοτήτων στον τομέα της υγείας κυρίως στο γηραιότερο πληθυσμό. Η συνεχιζόμενη άνοδος του προσδόκιμου χρόνου ζωής σε συνδυασμό με την υπογεννητικότητα αυξάνει τον αριθμό των ηλικιωμένων ατόμων. Η κοινωνικοοικονομική κατάσταση των ηλικιωμένων ατόμων εξαρτάται από τη νοσηρότητα (υποκειμενική και αντικειμενική) καθώς και από το εισόδημα (ατομικό εισόδημα και εισόδημα νοικοκυριού).

Χρησιμοποιώντας κάποιους δείκτες, το πρόγραμμα SHARE μπορεί να ελέγξει και να συγκρίνει τους παράγοντες που επηρεάζουν τις κοινωνικοοικονομικές ανισότητες μεταξύ των Ευρωπαϊκών χωρών. Η παρούσα εργασία θα ασχοληθεί με την εύρεση αυτών των δεικτών και την ανάλυσή τους μεταξύ των χωρών του 2^{ου} κύματος.

Η εργασία στοχεύει στην εφαρμογή καθιερωμένων μεθόδων όπως εκτίμηση κοινωνικοοικονομικών ανισοτήτων και στατιστική ανάλυση με σκοπό τη διερεύνηση κοινωνικοοικονομικών ανισοτήτων μεταξύ ατόμων ηλικίας 50 ετών και άνω σε χώρες της Ευρωπαϊκής Ένωσης. Για τη στατιστική ανάλυση χρησιμοποιούνται οικονομικές μεταβλητές, μεταβλητές υγείας καθώς και δημογραφικές μεταβλητές. Αποτελείται κυρίως από πέντε κεφάλαια. Αρχικά γίνεται μια ιστορική αναφορά στο ερευνητικό πρόγραμμα SHARE, συνεχίζοντας στο 2^ο κεφάλαιο με τη θεωρητική προσέγγιση της εργασίας. Πιο συγκεκριμένα, αναφέρονται οι μέθοδοι εκτίμησης των κοινωνικοοικονομικών ανισοτήτων εστιάζοντας στην Καμπύλη Lorenz, στο Δείκτη Συγκέντρωσης και στο εύρος των πιθανών τιμών του καθώς και η Λογιστική Παλινδρόμηση. Τα δύο τελευταία μεθοδολογικά πλαίσια αποτελούν τα κύρια ζητήματα της εργασίας. Ακόμα, καθορίζονται οι τρόποι επιλογής κατάλληλου μοντέλου μέσω κάποιων test. Ακολουθεί το 3^ο κεφάλαιο όπου εισάγονται οι μεταβλητές και γίνεται μια πρώτη περιγραφική ανάλυση των δεδομένων. Στη συνέχεια, εφαρμόζεται ο Δείκτης Συγκέντρωσης καθώς και η Λογιστική Παλινδρόμηση στα δεδομένα της εργασίας για τέσσερις μεταβλητές υγείας και τέλος δίνονται κάποια γενικά συμπεράσματα για τις δημογραφικές, οικονομικές και μεταβλητές υγείας.

2. ΘΕΩΡΗΤΙΚΟ ΥΠΟΒΑΘΡΟ

2.1 ΤΟ ΕΡΕΥΝΗΤΙΚΟ ΠΡΟΓΡΑΜΜΑ SHARE

Στην παρούσα εργασία τα δεδομένα που χρησιμοποιούνται προέρχονται από το Ερευνητικό Πρόγραμμα SHARE (Survey of Health Ageing and Retirement in Europe). Στο συγκεκριμένο κεφάλαιο θα αναφερθούμε πιο αναλυτικά στο πρόγραμμα αυτό, στη χρησιμότητά του καθώς και στο ιστορικό της δημιουργίας του.¹

Το ερευνητικό πρόγραμμα SHARE αποτελεί μια πρωτοποριακή προσέγγιση και εξαιρετικά χρήσιμη μελέτη για την υγεία, τη γήρανση και τη συνταξιοδότηση στις χώρες της Ευρώπης για άτομα ηλικίας 50 ετών και άνω. Κύριος στόχος του είναι η μελέτη της διαδικασίας μετάβασης από την ενεργό ζωή στη συνταξιοδότηση. Πώς όμως δημιουργήθηκε η ιδέα ενός τέτοιου προγράμματος; Η αρχική ιδέα του προγράμματος SHARE βασίζεται κυρίως σε δύο άλλα ερευνητικά πρότυπα. Ο συνδυασμός μιας διαχρονικής προσέγγισης με διεπιστημονικό σχεδιασμό έγινε αρχικά από το Βρετανικό ερευνητικό πρόγραμμα ELSA (English Longitudinal Survey on Ageing) και την Αμερικανική έρευνα HRS (US Health and Retirement Study).

Η διαχρονική έρευνα για τη γήρανση στη Βρετανία (ELSA) συλλέγει δεδομένα από ένα αντιπροσωπευτικό δείγμα Βρετανικού πληθυσμού ατόμων ηλικίας 50 ετών και άνω. Η συλλογή δεδομένων ξεκίνησε το 1998 και μετρά ήδη 7 κύματα ολοκληρωμένων δεδομένων, με το 8^ο κύμα να ολοκληρώνεται αυτή την περίοδο (Μάιος 2016-Ιούνιος 2017). Τα δεδομένα έχουν σχεδιαστεί για να χρησιμοποιηθούν στη μελέτη ενός ευρύτερου συνόλου θεμάτων που σχετίζονται με την κατανόηση της διαδικασίας της γήρανσης καλύπτοντας θέματα όπως τους δεσμούς μεταξύ οικονομικής θέσης, υγείας, εκπαιδευτικού επιπέδου και ψυχικής υγείας, το προσδόκιμο υγιούς ζωής κ.α..

Αντίστοιχα, η μελέτη για την υγεία και τη συνταξιοδότηση στην Αμερική (HRS) αποτελεί και αυτή μια διαχρονική έρευνα για ένα αντιπροσωπευτικό δείγμα Αμερικανών ηλικίας 50 ετών και άνω. Η κύρια ιδέα αυτής της μελέτης δημιουργήθηκε με σκοπό να εξεταστούν οι τρόποι αλληλεπίδρασης των ηλικιωμένων ατόμων με τους κοινωνικούς, τους οικονομικούς και τους ψυχολογικούς παράγοντες καθώς και οι αποφάσεις συνταξιοδότησης προς ενημέρωση της εθνικής συζήτησης συνταξιοδότησης όσον αφορά το γηραιό πληθυσμό. Η πρώτη συλλογή δεδομένων ξεκίνησε το 1992 και έκτοτε εφαρμόζεται κάθε δύο χρόνια. Μερικά από τα ζητήματα της μελέτης είναι η υγειονομική περιθαλψη, η στέγαση, τα περιουσιακά στοιχεία, οι συντάξεις, η απασχόληση, η αναπηρία κ.α..

Τα δεδομένα των δύο προαναφερθέντων μελετών (ELSA, HRS) χρησιμοποιήθηκαν για να συγκριθούν άμεσα τα μέτρα υγείας, εκπαίδευσης και εισοδήματος μεταξύ ατόμων ηλικίας 55-64

¹ Πηγή: Κεντρική ιστοσελίδα του προγράμματος SHARE (www.share-project.org)

ετών στην Αγγλία και τις ΗΠΑ. Τα πρώτα αποτελέσματα διαπίστωσαν κάποιες διαφορές μεταξύ των δύο χωρών οι οποίες δεν οφείλονταν στο σχεδιασμό της μελέτης ή στις μετρήσεις. Πιο συγκεκριμένα, οι διαφορές στον τομέα της υγείας (ποσοστά νοσηρότητας, παράγοντες κινδύνου παχυσαρκίας κ.λπ.) αποδείχτηκαν πραγματικές και όχι αποτέλεσμα διαφοροποιημένης συμπεριφοράς ανάμεσα στις δύο διαφορετικές αυτές χώρες. Άλλες παρόμοιες μελέτες είναι η Ιρλανδική διαχρονική μελέτη για τη γήρανση (TILDA), η Έρευνα Κινέζικης Υγείας και Συνταξιοδότησης (CHARLS), η διαχρονική μελέτη της υγείας των ηλικιωμένων της Βραζιλίας (ELSI-Brazil), η Ιαπωνική μελέτη γήρανσης και συνταξιοδότησης (JSTAR), η Κορεατική διαχρονική μελέτη της γήρανσης (KLOSA), η διαχρονική μελέτη γήρανσης στην Ινδία (LASI) καθώς και η Μεξικανική μελέτη για την υγεία και τη γήρανση (MHAS).

Έχοντας ως προσχέδιο τις δύο πρώτες “αδερφικές” μελέτες, το 2002 εδραιώθηκε η δημιουργία της βάσης δεδομένων της μελέτης SHARE η οποία, όπως προαναφέραμε, ερευνά για την υγεία, τη γήρανση και τη συνταξιοδότηση στην Ευρώπη για άτομα ηλικίας 50 ετών και άνω. Το πρόγραμμα αυτό ανταποκρίθηκε στην ανακοίνωση της Ευρωπαϊκής Επιτροπής για την εξέταση της δυνατότητας της δημιουργίας, σε συνεργασία των κρατών μελών, μιας Ευρωπαϊκής έρευνας για τη μακροχρόνια γήρανση.

Συντονίζεται στο Μόναχο της Γερμανίας από τον Axel Börsch-Supan, στο Κέντρο Μονάδων Οικονομίας της Γήρανσης (Munich Center of Economics of Aging, MEA), ένα τμήμα του Ερευνητικού Ινστιτούτου Max Planck για το Κοινωνικό Δίκτυο και την Κοινωνική Πολιτική (Max Planck Institute for Social Law and Social Policy, MPISOC) καθώς και στην Ιταλία και στην Ολλανδία. Αποτελείται από ένα επιστημονικό συμβούλιο με περισσότερους από 150 εξέχοντες διεθνείς ερευνητές παγκοσμίως που χωρίζονται σε διεπιστημονικές εθνικές ομάδες, οι οποίες συμβάλουν στη διατήρηση και στη βελτίωση των υψηλών επιστημονικών προτύπων του προγράμματος.

Ένα από τα πλεονεκτήματά του προγράμματος SHARE σε σχέση με τις μελέτες ELSA και HRS είναι ότι βασίζεται στο σχεδιασμό ενός πίνακα ο οποίος καταγράφει τα χαρακτηριστικά της διαδικασίας γήρανσης δίνοντας μια πλήρη εικόνα. Τα στοιχεία που συλλέγονται περιλαμβάνουν μεταβλητές υγείας, ψυχολογικές και οικονομικές μεταβλητές, μεταβλητές κοινωνικής υποστήριξης καθώς και βιοδείκτες (π.χ. δείκτης μάζας σώματος). Οι μεταβλητές και οι δείκτες δημιουργούνται από το πρόγραμμα AMANDA RTD και τα δεδομένα είναι διαθέσιμα δωρεάν σε ολόκληρη την ερευνητική κοινότητα.

Μέχρι τώρα έχουν οργανωθεί 6 κύματα για τη συλλογή πληροφοριών από άτομα ηλικίας 50 ετών και άνω σε διάφορες Ευρωπαϊκές χώρες. Η χρηματοδότησή τους γίνεται κυρίως από την Ευρωπαϊκή Επιτροπή, το Εθνικό Ινστιτούτο των ΗΠΑ για τη γήρανση καθώς και από το Γερμανικό Ομοσπονδιακό Υπουργείο Παιδείας και Έρευνας.

Το πρώτο κύμα του προγράμματος SHARE πραγματοποιήθηκε το 2004 με τη συμβολή 11 Ευρωπαϊκών χωρών, από τη Σκανδιναβία έως την Κεντρική Ευρώπη και τη Μεσόγειο. Το Ισραήλ εντάχθηκε στο πλαίσιο της μελέτης στα τέλη του 2004 και αποτελεί την πρώτη χώρα στη Μέση

Ανατολή που ξεκίνησε μια συστηματική μελέτη του γηραιού πληθυσμού. Το ερωτηματολόγιο του πρώτου κύματος περιείχε 20 ενότητες που αφορούσαν την υγεία, την οικονομική κατάσταση και τα κοινωνικά δίκτυα του ατόμου. Τα δεδομένα έχουν συλλεχθεί από συνεντεύξεις προσωπικές (πρόσωπο με πρόσωπο), με τη βοήθεια ηλεκτρονικού υπολογιστή κι ενός προγράμματος φιλικού προς το χρήστη, το πρόγραμμα CAPI (Computer Assisted Personal Interviewing) και συμπληρώνοντας ένα αυτοματοποιημένο ερωτηματολόγιο (ερωτήσεις πολλαπλής επιλογής).

Το δεύτερο κύμα διεξήχθη το 2006-2007 με την εισχώρηση τριών κρατών της Ευρωπαϊκής Ένωσης, της Τσέχικης Δημοκρατίας, της Πολωνίας και της Ιρλανδίας. Η καινοτομία του κύματος αυτού ήταν η ανάδειξη ενός “επαναληπτικού” δείγματος ερωτηθέντων, πέραν αυτών του πρώτου κύματος, σε όλες τις χώρες που πήραν μέρος στο πρώτο κύμα με εξαίρεση την Αυστρία και του Φλαμανδικού τμήματος του Βελγίου. Για το δείγμα αυτό χρησιμοποιήθηκαν οι ίδιες μέθοδοι δειγματοληψίας, η συλλογή δεδομένων έγινε με τον ίδιο τρόπο όπως στο πρώτο κύμα και το ερωτηματολόγιο αποτελούνταν από 22 ενότητες. Ένα ακόμα ιδιαίτερο χαρακτηριστικό αυτού του κύματος ήταν ότι πραγματοποιήθηκαν συνεντεύξεις για το τέλος της ζωής των νεκρών ερωτηθέντων του πρώτου κύματος.

Το τρίτο κύμα συλλογής δεδομένων έλαβε χώρα το 2008-2009 και επικεντρώθηκε περισσότερο στο ιστορικό της ζωής των ερωτηθέντων γι αυτό και ονομάστηκε SHARELIFE. Στο κύμα αυτό συλλέχθηκαν μεμονωμένα στοιχεία από το σύνολο της ζωής των ερωτηθέντων και θεσμικά μακροοικονομικά δεδομένα για το κράτος πρόνοιας με στόχο την αξιολόγηση της επίδρασης των παρεμβάσεων της κοινωνικής πρόνοιας στη ζωή του ατόμου. Άμεση συνέπεια του κύματος αυτού είναι μια πρώτη προσπάθεια αξιολόγησης των πολιτικών σε όλη την Ευρώπη. Τα δεδομένα του SHARELIFE είναι ιδανικά για την έρευνα στους τομείς της κοινωνιολογίας, της οικονομίας, της γεροντολογίας και της δημογραφίας χάρη στην ποικιλομορφία του ερωτηματολογίου όσον αφορά σημαντικούς τομείς της ζωής των ερωτηθέντων.

Το τέταρτο κύμα πραγματοποιήθηκε το 2010-2011 με την εισαγωγή της Εσθονίας, της Ουγγαρίας, της Πορτογαλίας και της Σλοβενίας στο πρόγραμμα. Μια πρωτοποριακή ενότητα κοινωνικού δικτύου συμπεριλήφθηκε στο ερωτηματολόγιο, βασισμένη σε μια προσέγγιση δημιουργίας ονόματος. Η καινοτομία αυτή κατέστησε το πρόγραμμα SHARE την πρώτη διεθνή έρευνα για την παροχή πληροφοριών κοινωνικού δικτύου με λεπτομερή και συγκριτικό τρόπο.

Το πέμπτο κύμα ολοκληρώθηκε το 2013 και περιλάμβανε νέες ενότητες στο ερωτηματολόγιο σχετικά με τις συνθήκες παιδικής ηλικίας, τη χρήση του υπολογιστή στην εργασία και τον κοινωνικό αποκλεισμό. Έλαβαν συμμετοχή 15 χώρες, μέσα στις οποίες και το Ισραήλ. Το έκτο κύμα διεξήχθη το 2015, όπου για δεύτερη φορά μετά το τέταρτο κύμα, το ερωτηματολόγιο περιείχε στοιχεία σχετικά με το κοινωνικό δίκτυο των ερωτηθέντων. Για πρώτη φορά στο πρόγραμμα συμμετείχε και η Κροατία. Τέλος, το έβδομο κύμα έχει ξεκινήσει να συλλέγεται από τις αρχές του 2017.

Μέχρι στιγμής το πρόγραμμα SHARE καλύπτει 27 Ευρωπαϊκές χώρες και το Ισραήλ. Έχει στο σύνολό του δεδομένα για περισσότερα από 120.000 άτομα ηλικίας 50 ετών και άνω, έχοντας

πραγματοποιήσει πάνω από 297.000 συνεντεύξεις. Τα στοιχεία αυτά είναι συγκρίσιμα μεταξύ των χωρών και των κυμάτων καθώς χρησιμοποιήθηκαν από όλα τα συμμετέχοντα κράτη τυποποιημένα και σταθμισμένα ερωτηματολόγια το οποία όχι μόνο περιλαμβάνουν τις ίδιες πληροφορίες (ερωτήσεις και μεταβλητές) αλλά χρησιμοποιούνται τα ίδια εργαλεία (λογισμικά) και οι ίδιες τεχνικές συμπλήρωσης ελλειπουσών παρατηρήσεων (imputations of missing values).

Η κάθε μια συμμετέχουσα χώρα είναι υπεύθυνη για τη δειγματοληψία και τη συλλογή των δεδομένων. Περισσότερες λεπτομέρειες και διαδικασίες καθώς και βάσεις δεδομένων βρίσκονται στην επίσημη κεντρική ιστοσελίδα του προγράμματος www.share-project.org.

2.2 ΜΕΘΟΔΟΙ ΕΚΤΙΜΗΣΗΣ ΚΟΙΝΩΝΙΚΟΟΙΚΟΝΟΜΙΚΩΝ ΑΝΙΣΟΤΗΤΩΝ

Υπάρχουν μέτρα τα οποία χρησιμοποιούνται ευρέως για τον προσδιορισμό των ανισοτήτων στον τομέα της υγείας. Τα κυριότερα μέτρα βάση της βιβλιογραφίας είναι το εύρος, η καμπύλη Lorenz και ο συντελεστής Gini, η κλίση και οι σχετικοί δείκτες ανισότητας καθώς και ο δείκτης συγκέντρωσης. Στην εργασία αυτή θα εστιάσουμε στο δείκτη συγκέντρωσης CI -ή C- (Concentration Index).

2.2.1 ΚΑΜΠΥΛΗ LORENZ

Η καμπύλη συγκέντρωσης των παθήσεων $L(s)$ οριοθετεί την αθροιστική αναλογία της μεταβλητής υγείας έναντι της αθροιστικής αναλογίας του πληθυσμού, ο οποίος κατατάσσεται βάσει κάποιās κοινωνικοοικονομικής κατάστασης, ξεκινώντας από το λιγότερο πλεονεκτικό άτομο έναντι της αθροιστικής αναλογίας της νόσου.

Αρχικά, το ενδιαφερόμενο δείγμα κατατάσσεται με βάση την κοινωνικοοικονομική του κατάσταση. Στην περίπτωση που η μεταβλητή κατάταξης είναι το εισόδημα των ατόμων του δείγματος, η κατάταξη ξεκινάει με το φτωχότερο άτομο και συνεχίζει μέσω της κατανομής του εισοδήματος έως το πλουσιότερο άτομο. Αυτή η σχετική κατάταξη του εισοδήματος απεικονίζεται γραφικά στον οριζόντιο άξονα ενώ στον κάθετο άξονα απεικονίζεται η αθροιστική αναλογία της μεταβλητής ασθένειας που μας ενδιαφέρει. Αυτό βέβαια προϋποθέτει ότι υπάρχει ένα βασικό μέτρο ασθένειας το οποίο μπορεί να καταμετρηθεί και να συγκριθεί σε ατομικό επίπεδο.

Η διαγώνια γραμμή (η γραμμή των 45 μοιρών) σημειώνεται ως η γραμμή της τέλει ισότητας όπου η αναλογία του πληθυσμού της νόσου είναι ανάλογη με το εισόδημα έτσι ώστε το φτωχότερο 20% των ατόμων να βιώνει το 20% της νόσου στο δείγμα.

2.2.2 ΔΕΙΚΤΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ

Ο δείκτης συγκέντρωσης (CI ή C) προέρχεται από την καμπύλη συγκέντρωσης L(s). Ορίζεται ως η μονάδα μείον το διπλάσιο της περιοχής μεταξύ της καμπύλης συγκέντρωσης L(s) και της γραμμής 45 μοιρών. Στην περίπτωση που δεν υπάρχει κοινωνικοοικονομική ανισότητα, ο δείκτης συγκέντρωσης είναι ίσος με μηδέν. Όταν η καμπύλη συγκέντρωσης είναι πάνω από τη διαγώνιο γραμμή τότε ο δείκτης συγκέντρωσης παίρνει αρνητική τιμή, υποδεικνύοντας δυσανάλογη συγκέντρωση της μεταβλητής υγείας μεταξύ των φτωχότερων ατόμων ευνοώντας τα περισσότερα πλεονεκτικά μέλη της κοινωνίας. Αντίστοιχα όταν η καμπύλη συγκέντρωσης είναι κάτω από τη διαγώνιο γραμμή τότε ο δείκτης συγκέντρωσης παίρνει θετική τιμή ευνοώντας τα λιγότερο πλεονεκτικά μέλη της κοινωνίας.

Ο δείκτης συγκέντρωσης ορίζεται ως εξής:

$$C = 1 - 2 \int_0^1 L(s) ds \quad (1)$$

Για μια διακριτή μεταβλητή, ο δείκτης συγκέντρωσης μπορεί να υπολογιστεί πιο απλά: (Kakwani et al., 1997)

$$C = \frac{2}{n \mu} \left(\sum_{i=1}^n x_i R_i \right) - 1 \quad (2)$$

όπου x_i είναι η μεταβλητή του τομέα της υγείας του i -οστού ατόμου, για $i=1, \dots, n$

μ είναι ο μέσος όρος της μεταβλητής υγείας, δηλαδή

$$\mu = \frac{1}{n} \sum_{i=1}^n x_i$$

και $R_i = i/n$ είναι η κλασματική τάξη του i -οστού ατόμου στην κατανομή του βιοτικού του επιπέδου, με $i = 1$ για τους φτωχότερους και $i = n$ για τους πλουσιότερους.

Η παραπάνω εξίσωση (2) καθιστά σαφή την εξάρτηση του δείκτη συγκέντρωσης C από την κοινωνικοοικονομική διάσταση στην κατανομή της μεταβλητής υγείας.² Γι αυτό το λόγο, ο δείκτης συγκέντρωσης C αποτελεί ένα ευρέως χρησιμοποιούμενο μέτρο για την ανίχνευση ανισοτήτων που σχετίζονται με το εισόδημα στον τομέα της υγείας.³

² *The Concentration Index, Chapter 8*

³ *Adam Wagstaff, Eddy van Doorslaer, Naoko Watanabe, Journal of Econometrics 112, 2003, 207-223*

2.2.3 ΕΥΡΟΣ ΠΙΘΑΝΩΝ ΤΙΜΩΝ ΔΕΙΚΤΗ ΣΥΓΚΕΝΤΡΩΣΗΣ

Όταν διερευνάται η ανισότητα μιας μεταβλητής του τομέα υγείας της οποίας η κατανομή δεν έχει όρια, ο δείκτης συγκέντρωσης βρίσκεται στο διάστημα $[-1,1]$. Ωστόσο, όταν η υπό εξέταση μεταβλητή είναι δίτιμη τότε ο μέσος όρος της κατανομής θέτει όρια στις πιθανές τιμές του. Πιο συγκεκριμένα, καθώς αυξάνεται η μέση τιμή, το εύρος των πιθανών τιμών που μπορεί να πάρει ο δείκτης συγκέντρωσης μικραίνει.⁴

Πρόταση: Έστω μ είναι η μέση τιμή μιας δίτιμης μεταβλητής x της οποίας η κατανομή με κάποιο μέτρο κοινωνικοοικονομικής κατάστασης είναι το αντικείμενο ενδιαφέροντος και n το μέγεθος του δείγματος. Δοθέντος $\mu \in (0,1)$, η ελάχιστη τιμή του δείκτη συγκέντρωσης είναι ίση με $\mu - 1 + \frac{1}{n}$ και η μέγιστη τιμή είναι ίση με $1 - \mu + \frac{1}{n}$.

Η απόδειξη είναι σχετικά απλή. Όπως προαναφέραμε, ο δείκτης συγκέντρωσης για δίτιμα δεδομένα γράφεται ως

$$C = \frac{2}{n\mu} \left(\sum_{i=1}^n x_i R_i \right) - 1 \quad (3)$$

Δοθέντος $\mu \in (0,1)$, το μέγιστο του δείκτη συγκέντρωσης είναι όταν το φτωχότερα άτομα j έχουν τιμή $x = 0$ και τα πλουσιότερα $n - j$ άτομα έχουν τιμή $x = 1$. Σε αυτή την περίπτωση ισχύει ότι

$$\mu = \frac{n-j}{n} \quad (4)$$

Οπότε έχουμε

$$\begin{aligned} C &= \frac{2}{n-j} \left[0 + 0 + \dots + \frac{j+1}{n} + \frac{j+2}{n} + \dots + 1 \right] - 1 \\ &= \frac{2}{n-j} \frac{1}{n} \left[0 + 0 + \dots + (j+1) + (j+2) + \dots + n \right] - 1 \\ &= \frac{2}{n-j} \frac{1}{n} \frac{[2(j+1) + (n-j-1)](n-j)}{2} - 1 \\ &= \frac{j+1+n}{n} - 1 \\ &\stackrel{(2)}{\implies} 1 - \mu + \frac{1}{n} \end{aligned}$$

⁴ Adam Wagstaff, "The bounds of the concentration index when the variable interest is binary, with an application to immunization inequality", *Health Economics* 14:429-432, 2005

Για μεγάλα δείγματα ο όρος $\frac{1}{n}$ τείνει στο 0 και η μέγιστη τιμή του δείκτη συγκέντρωσης τείνει στο 1 - μ. Όμοια ορίζεται και η ελάχιστη τιμή του δείκτη συγκέντρωσης.

2.3 ΜΕΘΟΔΟΛΟΓΙΚΟ ΠΛΑΙΣΙΟ

Η πιο συνηθισμένη μορφή κατηγορικών δεδομένων είναι τα δίτιμα δεδομένα (binary data). Τα δεδομένα αυτά διαχωρίζονται από τα ομαδοποιημένα καθώς γνωρίζουμε για κάθε άτομο μέσα στο δείγμα την τιμή απόκρισης Y (0 = αποτυχία, 1 = επιτυχία). Υπάρχουν τρεις συναρτήσεις σύνδεσης που χρησιμοποιούμε για δίτιμα δεδομένα: η Logit, η Probit και η Complementary log-log. Στη συγκεκριμένη εργασία χρησιμοποιούμε σα συνάρτηση σύνδεσης τη συνάρτηση Logit, οπότε μιλάμε για ένα μοντέλο λογιστικής παλινδρόμησης (Logistic Regression).

2.3.1 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Η πιο συνηθισμένη μέθοδος ανάλυσης δίτιμων δεδομένων στον τομέα των κοινωνικών ερευνών είναι η Λογιστική Παλινδρόμηση (Binary Logistic Regression). Χρησιμοποιείται κυρίως σε καταστάσεις πρόβλεψης της ύπαρξης ή απουσίας ενός χαρακτηριστικού ή ενός συμβάντος. Πιο συγκεκριμένα, η μεταβλητή απόκρισης είναι κατηγορική και δίτιμη, η οποία παίρνει τις τιμές 0 και 1 που συμβολίζουν αντίστοιχα την έλλειψη ή την ύπαρξη του χαρακτηριστικού ή του συμβάντος.

Η Λογιστική Παλινδρόμηση υπερέχει έναντι της απλής Παλινδρόμησης (OLS) όσον αφορά την ανάλυση δίτιμων δεδομένων καθώς η χρήση της επιτρέπει λιγότερο αυστηρές προϋποθέσεις. Πιο αναλυτικά, δεν χρειάζεται να ισχύουν οι υποθέσεις κανονικότητας και ομοσκεδαστικότητας των επεξηγηματικών μεταβλητών. Ακόμη, εκτιμά τις μεταβολές του λογαρίθμου των σχετικών πιθανοτήτων της μεταβλητής απόκρισης και όχι απλά τις μεταβολές στη μεταβλητή απόκριση, όπως εφαρμόζεται με την απλή γραμμική παλινδρόμηση. Τέλος, η Λογιστική Παλινδρόμηση χρησιμοποιείται επίσης και για την εκτίμηση της επίδρασης κάθε επεξηγηματικής μεταβλητής στη διαμόρφωση των τιμών της μεταβλητής απόκρισης.

Έστω λοιπόν ότι έχουμε μια δίτιμη μεταβλητή απόκρισης Y και μια επεξηγηματική μεταβλητή X. Τότε η πιθανότητα επιτυχίας εκφράζεται από τη συνάρτηση p(x), όπου

$$p(x) = \frac{e^{\alpha + \beta x}}{1 + e^{\alpha + \beta x}} \quad (5)$$

Επομένως, η σχετική πιθανότητα επιτυχίας (odds) ισοδυναμεί με

$$\frac{p(x)}{1 - p(x)} = e^{\alpha + \beta x} \quad (6)$$

Η συνάρτηση logit αναφέρεται στο λογάριθμο της σχετικής πιθανότητας του ενδεχομένου που μας ενδιαφέρει (της σχετικής πιθανότητας επιτυχίας) και δίνεται ως εξής

$$\text{logit} [p(x)] = \log \left[\frac{p(x)}{1-p(x)} \right] = \alpha + \underline{\beta}x \quad (7)$$

δηλαδή η συνάρτηση logit προκύπτει σα μια συνάρτηση που συνδέει την πιθανότητα επιτυχίας με την τιμή της ερμηνευτικής μεταβλητής.

Ακόμη, ο λόγος σχετικών πιθανοτήτων (odds ratio) ενός ενδεχομένου x_1 ως προς το ενδεχόμενο x_2 είναι ίσος με

$$\frac{P(Y = 1|X = x_1) / P(Y = 0|X = x_1)}{P(Y = 1|X = x_2) / P(Y = 0|X = x_2)} = \frac{P(x_1) / [1 - p(x_1)]}{P(x_2) / [1 - p(x_2)]} \quad (8)$$

όπου στη Λογιστική Παλινδρόμηση ισοδυναμεί με $e^{\underline{\beta}(x_1 - x_2)}$ και δείχνει πόσες φορές η σχετική πιθανότητα του ενδεχομένου x_1 είναι μεγαλύτερη από τη σχετική πιθανότητα του ενδεχομένου x_2 .

Οι λόγοι σχετικών πιθανοτήτων συνήθως ερμηνεύονται σε σχέση με ένα επίπεδο αναφοράς κάθε παράγοντα και η εκτίμηση των συντελεστών των επεξηγηματικών μεταβλητών γίνεται με βάση τη Μέθοδο της Μέγιστης Πιθανοφάνειας (Maximum Likelihood). Για τη σωστή εφαρμογή της Λογιστικής Παλινδρόμησης και την αξιοπιστία του αποτελέσματός της απαιτείται μεγάλο δείγμα δεδομένων.

2.3.2 ΠΡΟΣΑΡΜΟΓΗ ΜΟΝΤΕΛΟΥ

Αρχικά γίνεται έλεγχος της σημαντικότητας του μοντέλου. Η πιο συνηθισμένη μέθοδος στη λογιστική παλινδρόμηση είναι ο δείκτης -2LL (-2 Log Likelihood) ο οποίος ισούται με μείον δύο φορές το λογάριθμο της πιθανοφάνειας. Το βέλτιστο μοντέλο θα έχει πιθανοφάνεια ίση με 1, δηλαδή -2LL=0. Συνεπάγεται ότι όσο η τιμή του δείκτη -2LL τείνει στο μηδέν τόσο καλύτερο είναι το μοντέλο.

Έχουν προταθεί διάφορες μέθοδοι για τον έλεγχο προσαρμογής ενός λογαριθμικού μοντέλου για δίτιμα δεδομένα. Ο συνηθέστερος έλεγχος που χρησιμοποιείται είναι ο έλεγχος των Hosmer and Lemeshow.

Το τεστ των Hosmer and Lemeshow υπολογίζει ένα δείκτη για τη συνολική προσαρμογή του μοντέλου. Πιο συγκεκριμένα, ελέγχει την υπόθεση ότι υπάρχει διαφορά μεταξύ των εκτιμώμενων και των παρατηρούμενων τιμών. Επιθυμητή είναι η σημαντικότητα του μοντέλου να είναι μεγαλύτερη από 0.05 έτσι ώστε να έχουμε καλή προσαρμογή του μοντέλου στα δεδομένα. Για τον έλεγχο, αρχικά διατάσσονται οι παρατηρήσεις ανάλογα με την προβλεπόμενη πιθανότητα επιτυχίας. Έπειτα ομαδοποιούνται οι παρατηρήσεις σε g ομάδες (συνήθως $g = 10$), με ίσο περίπου

αριθμό παρατηρήσεων και για καθεμία από αυτές καταγράφεται ο αριθμός των επιτυχιών και των αποτυχιών σχηματίζοντας έναν πίνακα $g \times 2$. Η στατιστική συνάρτηση των Hosmer and Lemeshow, X_{HL}^2 , είναι το X^2 του Pearson για τον παραπάνω πίνακα. Κάτω από τη μηδενική υπόθεση ότι οι παρατηρηθείσες τιμές δε διαφέρουν από τις εκτιμώμενες, η συνάρτηση ακολουθεί κατανομή $\chi_g^2 - 2$. Η απόρριψη της μηδενικής υπόθεσης δηλώνει ότι το μοντέλο είναι ανεπαρκές για το συγκεκριμένο επίπεδο σημαντικότητας του ελέγχου. Ένα μειονέκτημα του ελέγχου αυτού είναι ο μεγάλος αριθμός παρατηρήσεων που χρειάζεται, τουλάχιστον 70, αφού για την εφαρμογή του χωρίζονται τα δεδομένα σε ισόποσες ομάδες.

Εναλλακτικά μια ένδειξη για το μέγεθος της μεταβλητότητας του δείγματος που τελικά ερμηνεύεται από την παλινδρόμηση δίνεται από το δείκτη Nagelkerke R^2 . Όσο αυξάνεται η τιμή του δείκτη είναι ένδειξη ότι κάθε μεταβλητή που εισάγεται στο μοντέλο, προσθέτει πληροφορία στην εξίσωση, με τη μέγιστη τιμή που μπορεί να πάρει να ισούται με 1.

Τέλος ο πίνακας ταξινόμησης δείχνει το ποσοστό των περιπτώσεων της εξαρτημένης μεταβλητής που προβλέπονται σωστά βάσει του μοντέλου. Οπότε όσο μεγαλύτερη είναι η τιμή του ποσοστού τόσο καλύτερη είναι η προσαρμογή του μοντέλου στα δεδομένα της μεταβλητής.

3. ΔΕΔΟΜΕΝΑ ΕΦΑΡΜΟΓΩΝ

3.1 ΜΕΤΑΒΛΗΤΕΣ

Τα δεδομένα της εργασίας προέρχονται από το δεύτερο κύμα του Ερευνητικού Προγράμματος SHARE (Survey of Health Ageing and Retirement in Europe) το οποίο διεξήχθη το 2006-2007. Ειδικότερα, το δείγμα που διαθέτουμε αποτελείται από 33851 δεδομένα τα οποία διακρίνονται στις εξής ακόλουθες 11 μεταβλητές:

Πίνακας 3.1 Πίνακας Μεταβλητών

mergeid	person identifier (fix across modules and waves)	προσωπικός κωδικός αναγνώρισης
country	country identifier	χώρα προέλευσης
chronic2w2	2+ chronic diseases (w2 version)	χρόνιες ασθένειες πάνω από 2
eurodcat	EURO-D caseness	EURO-D περίπτωση
gali	limitations with activities-gali	περιορισμός στις δραστηριότητες
gender	gender	φύλο
age	age	ηλικία
yedu	years of education	χρόνια εκπαίδευσης
sphus	self-perceived health -US scale	αυτοεκτίμηση υγείας -Αμερικάνικη κλίμακα
hhnetw	household net worth average imputed	μέση καθαρή αξία τεκμαρτών των νοικοκυριών
hhinc	household income version A	εισόδημα του νοικοκυριού (έκδοση A)

Μεταβλητή mergeid

Η μεταβλητή mergeid αποτελεί μια βασική μεταβλητή για την ομαλή συγχώνευση διαφορετικών ενοτήτων και/ή κυμάτων σε ατομικό επίπεδο. Περιλαμβάνεται σε όλες τις ενότητες που περιέχουν τις απαντήσεις των ατόμων και γι αυτό το λόγο μπορούν να χρησιμοποιηθούν για τον συνδυασμό διαφορετικών ενοτήτων σε ατομικό επίπεδο. Ακόμα, είναι ένα μοναδικό και αμετάβλητο αναγνωριστικό για κάθε άτομο σε όλα τα κύματα και έχει τη μορφή “CC-hhhhhh-IT”, όπου το “CC” αναφέρεται στο σύντομο κωδικό της χώρας, το “hhhhhh” είναι ψηφία που προσδιορίζουν το νοικοκυριό και το “IT” είναι ο αναγνωριστικός κωδικός του ερωτώμενου μέσα σε κάθε νοικοκυριό. Πιο συγκεκριμένα, η κωδικοποίηση του “CC” παρουσιάζεται στον Πίνακα 6.1 στο παράρτημα.

Μεταβλητή country

Η μεταβλητή country είναι μια κατηγορική, μη διατάξιμη μεταβλητή (nominal) η οποία αναφέρεται στη χώρα προέλευσης του ερωτώμενου. Στο δείγμα του δεύτερου κύματος που χρησιμοποιούμε στην παρούσα εργασία τα δεδομένα προέρχονται από 14 χώρες και η κωδικοποίηση δίνεται στον Πίνακα 6.2 στο παράρτημα.

Μεταβλητή chronic2w2

Η μεταβλητή chronic2w2 δηλώνει την ύπαρξη ή μη περισσότερων από 2 χρόνιων ασθενειών. Είναι μια μεταβλητή η οποία δημιουργήθηκε από ένα σύνολο άλλων μεταβλητών (generated variable). Επίσης, είναι κατηγορική, μη διατάξιμη μεταβλητή (nominal) και η κωδικοποίηση περιλαμβάνεται στον Πίνακα 6.3 στο παράρτημα.

Μεταβλητή eurodcat

Η μεταβλητή eurod αναπτύχθηκε σε μια προσπάθεια δημιουργίας μιας κοινής κλίμακας μέτρησης συμπτωμάτων κατάθλιψης. Μαζί με τη μεταβλητή eurodcat αποτελούν δύο παραγόμενες μεταβλητές της ενότητας gv_health.

Πιο συγκεκριμένα, η eurod παράγεται από ερωτήσεις του τομέα ψυχικής υγείας και ειδικότερα είναι ένας σύνθετος δείκτης 16 στοιχείων (mh002-mh017)⁵. Η δημιουργία του δείκτη έχει ως σκοπό τη μείωση των 16 παραπάνω στοιχείων σε 12. Τα στοιχεία αυτά παρουσιάζονται στο παράρτημα στον Πίνακα 6.4.

Η μεταβλητή eurod αποτελεί το άθροισμα των παραπάνω 12 στοιχείων. Με αυτό τον τρόπο, το μέγιστο σκορ που μπορεί να έχει ένας ερωτώμενος είναι 12 που εξηγείται ως “πολύ καταθλιπτικός” και το ελάχιστο σκορ είναι 0 που εξηγείται ως “καθόλου καταθλιπτικός”.

Η επίτευξη του βαθμολογικού σκορ 4 και πάνω κατηγοριοποιείται ως “περίπτωση κατάθλιψης” και 3 και κάτω ως “περίπτωση μη κατάθλιψης”. Με αυτό τον τρόπο η παραγόμενη μεταβλητή eurodcat είναι μια είναι κατηγορική, μη διατάξιμη δίτιμη μεταβλητή (nominal, binary variable) και κωδικοποιείται με 1- “περίπτωση κατάθλιψης” και με 0- “περίπτωση μη κατάθλιψης”.

Μεταβλητή gali

Η μεταβλητή gali αναφέρεται στον περιορισμό των δραστηριοτήτων του ερωτηθέντος. Έχει αναπτυχθεί ως δείκτης για τη σύγκριση του προσδόκιμου επιβίωσης και της αναπηρίας σε

⁵ Ο πίνακας με τα 16 αυτά στοιχεία βρίσκεται στο τέλος του κεφαλαίου.

ολόκληρη την Ευρώπη. Πρόκειται για ένα παγκόσμιο ενιαίο όργανο μέτρησης μακροχρόνιων περιορισμών δραστηριοτήτων (6 ή περισσότερους μήνες) οι οποίοι αναφέρονται σε γενικά προβλήματα υγείας καθώς και στις δραστηριότητες που συνήθως κάνουν οι ερωτώμενοι.

Η μεταβλητή *galí*, όπως και η *euurodcat*, αποτελεί παραγόμενη μεταβλητή της ενότητας *gn_health*. Είναι μια είναι κατηγορική, μη διατάξιμη δίτιμη μεταβλητή (nominal, binary variable) και κωδικοποιείται με 1- “περιορισμένος” και με 0- “μη περιορισμένος”.

Μεταβλητή gender

Η μεταβλητή *gender* δηλώνει το φύλο του ερωτηθέντος. Είναι μια είναι κατηγορική, δίτιμη μεταβλητή (binary) και η κωδικοποίηση δίνεται στον Πίνακα 6.5 στο παράρτημα.

Μεταβλητή age

Η μεταβλητή *age* είναι μια αριθμητική μεταβλητή (scale) η οποία δηλώνει την ηλικία του ερωτηθέντος. Στις περιπτώσεις που δε συμπληρώθηκε η συγκεκριμένη μεταβλητή (missing value) έχει οριστεί αυθαίρετα να τοποθετείται η τιμή -99. Για καλύτερη στατιστική ανάλυση δημιουργήσαμε τη μεταβλητή *ageclass* η οποία είναι μια κατηγορική, διατάξιμη μεταβλητή (ordinal) που κατηγοριοποιεί τις ηλικίες των ερωτηθέντων σε 11 κλάσεις, οι οποίες δίνονται στον Πίνακα 6.6 στο παράρτημα, καθώς και τη μεταβλητή *agegroup* η οποία είναι μια κατηγορική, δίτιμη μεταβλητή που κωδικοποιείται με 1- “άτομα ηλικίας 65 ετών και άνω” και με 0- “άτομα ηλικίας 50 έως 64”.

Μεταβλητή yedu

Η μεταβλητή *yedu* είναι μια αριθμητική μεταβλητή (scale) η οποία δηλώνει τα χρόνια εκπαίδευσης του ερωτηθέντος. Στις περιπτώσεις που δε συμπληρώθηκε η συγκεκριμένη μεταβλητή (missing value) έχει οριστεί αυθαίρετα να τοποθετείται η τιμή -99. Για καλύτερη στατιστική ανάλυση δημιουργήσαμε τη μεταβλητή *yedu13* η οποία είναι μια κατηγορική, δίτιμη μεταβλητή που κωδικοποιείται με 1- “χρόνια εκπαίδευσης 13 ετών και άνω” (τριτοβάθμια εκπαίδευση) και με 0- “χρόνια εκπαίδευσης έως 12 έτη” (πρωτοβάθμια-δευτεροβάθμια εκπαίδευση).

Μεταβλητή sphus

Η μεταβλητή *sphus* είναι μια είναι κατηγορική, διατάξιμη μεταβλητή (ordinal) η οποία υπολογίζει την αυτό-αναφερόμενη υγεία και αποτελεί παραγόμενη μεταβλητή της ενότητας *gn_health*. Οι ερωτηθέντες βαθμολογούν την τρέχουσα γενική τους υγεία στην κλίμακα Likert,

Αμερικανικής έκδοσης των 5 βαθμών. Ο Πίνακας 6.7 στο παράρτημα δηλώνει τις κατηγορίες της μεταβλητής. Για καλύτερη στατιστική ανάλυση δημιουργήσαμε τη μεταβλητή `sphus_new` η οποία είναι μια κατηγορική, δίτιμη μεταβλητή που κωδικοποιείται με 1- “κακή αυτό-αναφερόμενη υγεία” και με 0- “καλή αυτό-αναφερόμενη υγεία”.

Μεταβλητή `hhnetw`

Η μεταβλητή `hhnetw` είναι μια αριθμητική μεταβλητή (scale) η οποία δηλώνει τη μέση καθαρή αξία τεκμαρτών του νοικοκυριού. Για καλύτερη στατιστική ανάλυση δημιουργήσαμε τη μεταβλητή `hhnetw_` η οποία είναι μια κατηγορική, δίτιμη μεταβλητή που κωδικοποιείται με 1- “μέση καθαρή αξία τεκμαρτών άνω της διαμέσου” και με 0- “μέση καθαρή αξία τεκμαρτών κάτω της διαμέσου”.

Μεταβλητή `hhinc`

Η μεταβλητή `hhinc` είναι μια αριθμητική μεταβλητή (scale) η οποία δηλώνει το εισόδημα του νοικοκυριού (έκδοση A). Για καλύτερη στατιστική ανάλυση δημιουργήσαμε τη μεταβλητή `hhinc_` η οποία είναι μια κατηγορική, δίτιμη μεταβλητή που κωδικοποιείται με 1- “εισόδημα νοικοκυριού άνω της διαμέσου” και με 0- “εισόδημα νοικοκυριού κάτω της διαμέσου”.

3.2 ΠΕΡΙΓΡΑΦΙΚΗ ΣΤΑΤΙΣΤΙΚΗ

Στην περιγραφική στατιστική γίνεται παρουσίαση των δεδομένων μέσω γραφικών και αριθμητικών μεθόδων. Ανάλογα με το είδος των μεταβλητών (ποιοτικές και ποσοτικές) υπάρχουν διάφοροι τρόποι παρουσίασης τους οποίους θα αναλύσουμε παρακάτω.

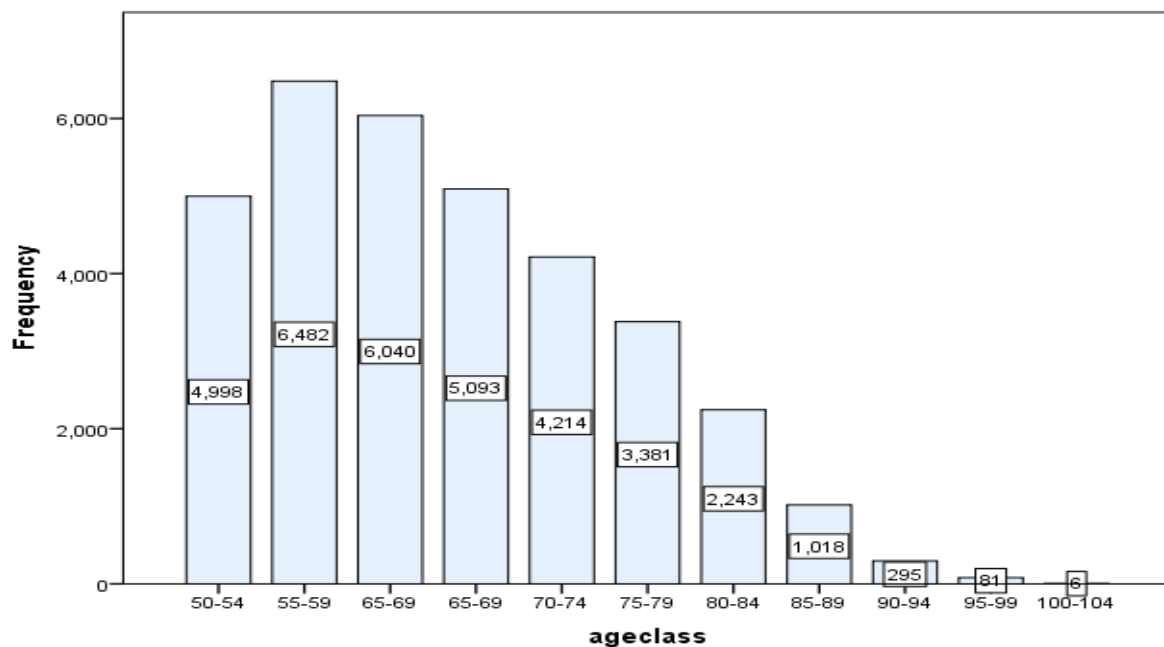
Όπως προαναφέραμε, το δείγμα των δεδομένων της παρούσας εργασίας αποτελείται από 33851 άτομα προερχόμενα από 14 Ευρωπαϊκές χώρες. Παρατηρούμε ότι οι χώρες με τα υψηλότερα ποσοστά συμμετοχής στην έρευνα είναι η Ελλάδα και το Βέλγιο (με ποσοστά 9.6% και 9.3% αντίστοιχα).

Πίνακας 3.2 Πίνακας συχνοτήτων μεταβλητής country

	Frequency	Percent	Valid Percent	Cumulative Percent
Αυστρία	1182	3.5	3.5	3.5
Γερμανία	2585	7.6	7.6	11.1
Σουηδία	2761	8.2	8.2	19.3
Ολλανδία	2638	7.8	7.8	27.1
Ισπανία	2382	7.0	7.0	34.1
Ιταλία	2931	8.7	8.7	42.8
Γαλλία	2879	8.5	8.5	51.3
Δανία	2548	7.5	7.5	58.8
Ελλάδα	3239	9.6	9.6	68.4
Ελβετία	1459	4.3	4.3	72.7
Βέλγιο	3143	9.3	9.3	82.0
Τσεχία	2671	7.9	7.9	89.9
Πολωνία	2426	7.2	7.2	97.0
Ιρλανδία	1007	3.0	3.0	100.0
Σύνολο	33851	100.0	100.0	

Το σύνολο των ατόμων είναι ηλικίας 50 έως 104 ετών με μέσο όρο ηλικίας τα 65.47 έτη όπως φαίνεται στο παρακάτω διάγραμμα κατανομής συχνοτήτων ηλικιακών ομάδων:

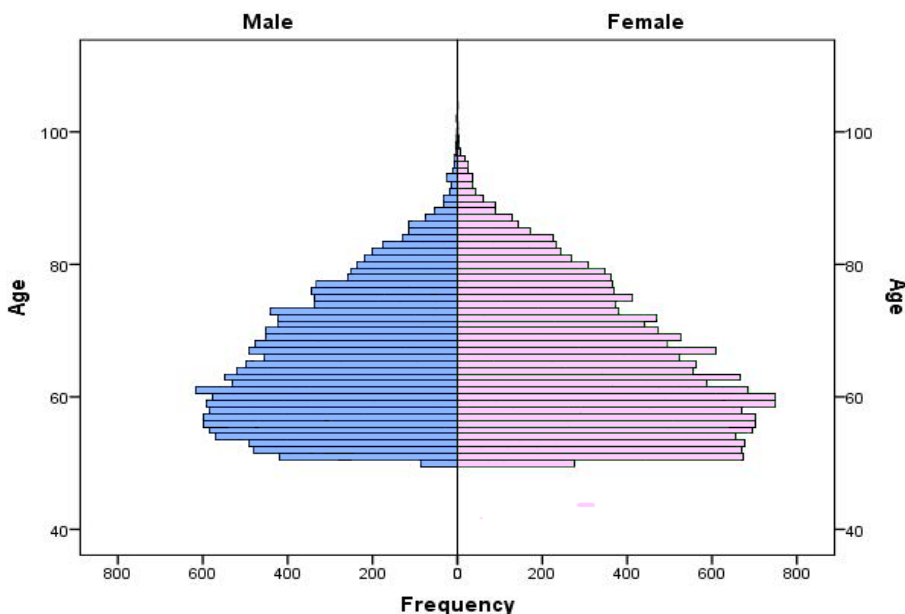
Διάγραμμα 3.1 Κατανομή συχνοτήτων ηλικιακών ομάδων



Από το σύνολο του πληθυσμού το 54.89% είναι γυναίκες και το 45.11% είναι άνδρες.

Ακολουθεί μια πυραμίδα ηλικιών για τα δύο φύλα ξεχωριστά. Σημαντική διαφορά παρατηρούμε κυρίως στις νεαρές ηλικίες των ερωτώμενων (ηλικίες 50-54 και 55-59) όπου οι γυναίκες πλεονεκτούν σε πληθυσμό σε σχέση με τους άνδρες.

Διάγραμμα 3.2 Πυραμίδα ηλικιών



Έπειτα παρουσιάζεται η μεταβλητή η οποία δηλώνει την ύπαρξη ή μη περισσότερων από 2 χρόνιων ασθενειών. Παρατηρούμε ότι το ποσοστό ελλειπουσών τιμών δεν είναι σημαντικό (μόλις 0.4%) και ότι παρ' όλο που τα ποσοστά είναι σχετικά κοντά, το μεγαλύτερο μέρος του πληθυσμού του δείγματός μας δηλώνει ότι έχει λιγότερο από 2 ασθένειες (με ποσοστό 55.5% έναντι 44.2%).

Πίνακας 3.3 Πίνακας συχνοτήτων μεταβλητής chronic2w2

	Frequency	Percent	Valid Percent	Cumulative Percent
λιγότερο από 2 ασθένειες	18772	55.5	55.7	55.7
2 ή περισσότερες ασθένειες	14946	44.2	44.3	100.0
Σύνολο	33718	99.6	100.0	
Missing System	133	0.4		
Total	33851	100.0		

Παρακάτω δίνεται ο πίνακας συχνοτήτων της μεταβλητής eurodcat η οποία κατατάσσει τους ερωτώμενους σε περιπτώσεις με κατάθλιψη ή μη. Πιο συγκεκριμένα, παρατηρούμε και σε αυτή την περίπτωση ότι το ποσοστό των ελλειπουσών τιμών είναι αρκετά μικρό (2.8%) και από τους ερωτώμενους που απάντησαν σε όλες τις ερωτήσεις που συμπεριελήφθησαν για τη διεξαγωγή του συμπεράσματος της εμφάνισης ή μη κατάθλιψης, ένα αρκετά μεγάλο ποσοστό ύψους 73.3% παρατηρήθηκε ότι δεν παρουσιάζει σημάδια κατάθλιψης έναντι ενός ποσοστού 23.9%.

Πίνακας 3.4 Πίνακας συχνοτήτων μεταβλητής eurodcat

	Frequency	Percent	Valid Percent	Cumulative Percent
Περιπτώσεις μη κατάθλιψης	24825	73.3	75.4	75.4
Περιπτώσεις κατάθλιψης	8092	23.9	24.6	100.0
Σύνολο	32917	97.2	100.0	
Missing System	934	2.8		
Total	33851	100.0		

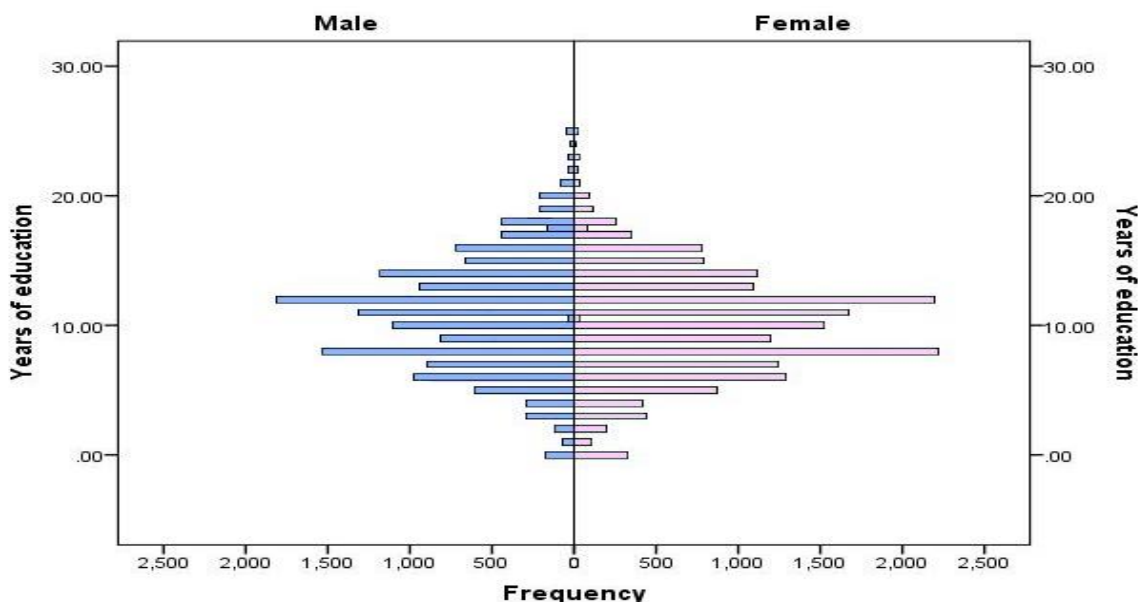
Ακόμη, ο παρακάτω πίνακας αναφέρεται στη μεταβλητή που δηλώνει αν ο ερωτώμενος παρουσιάζει κάποιους περιορισμούς όσον αφορά τις δραστηριότητές του. Ειδικότερα, παρατηρούμε κι εδώ ένα αρκετά μικρό ποσοστό ελλειπουσών τιμών (μόλις 0.3%) και από τους ερωτηθέντες που απάντησαν, το 56.6% δήλωσαν ότι τους τελευταίους 6 ή περισσότερους μήνες δεν παρουσίασαν κάποιον περιορισμό όσον αφορά τις δραστηριότητές τους.

Πίνακας 3.5 Πίνακας συχνοτήτων μεταβλητής gali

	Frequency	Percent	Valid Percent	Cumulative Percent
Μη περιορισμένος	19174	56.6	56.8	56.8
Περιορισμένος	14578	43.1	43.2	100.0
Σύνολο	33752	99.7	100.0	
Missing System	99	0.3		
Total	33851	100.0		

Ακολουθεί η πυραμίδα χρόνων εκπαίδευσης για τα δύο φύλα όπου αξίζει να αναφερθεί ότι περισσότερες γυναίκες δηλώνουν ότι έχουν ολοκληρώσει δευτεροβάθμια εκπαίδευση (12 έτη εκπαίδευσης) σε σχέση με τους άντρες.

Διάγραμμα 3.3 Πυραμίδα χρόνων εκπαίδευσης



Παρακάτω δίνεται ο πίνακας συχνοτήτων για τη μεταβλητή που υπολογίζει την αυτό-αναφερόμενη υγεία. Παρατηρούμε ότι δεν υπάρχουν ελλείπουσες τιμές και ότι το μεγαλύτερο ποσοστό του πληθυσμού δηλώνει πως έχει σχετικά καλή υγεία (37%).

Πίνακας 3.6 Πίνακας συχνοτήτων για τη μεταβλητή *sphus*

	Frequency	Percent	Valid Percent	Cumulative Percent
εξαιρετική	3032	9.0	9.0	9.0
πολύ καλή	6269	18.5	18.5	27.5
καλή	12527	37.0	37.0	64.5
μέτρια	8411	24.8	24.8	89.3
φτωχή	3612	10.7	10.7	100.0
Σύνολο	33851	100.0	100.0	

Τέλος παρουσιάζουμε στον παρακάτω πίνακα τα περιγραφικά μέτρα για δύο μεταβλητές οι οποίες ασχολούνται με την οικονομική κατάσταση του νοικοκυριού και πιο συγκεκριμένα αναφερόμαστε στη μέση καθαρή αξία τεκμαρτών και στο εισόδημα του νοικοκυριού των ερωτηθέντων. Οι μεταβλητές αυτές έχουν υψηλό ποσοστό ελλειπουσών τιμών (κοντά στο 20%) και έχουν εκτιμηθεί μερικώς, με χρήση στατιστικών μεθόδων (*imputation techniques*) από την ομάδα του προγράμματος SHARE⁶.

⁶ Paccagnella & Weber (2005)

Πίνακας 3.7 Περιγραφικά μέτρα των οικονομικών μεταβλητών του πληθυσμού

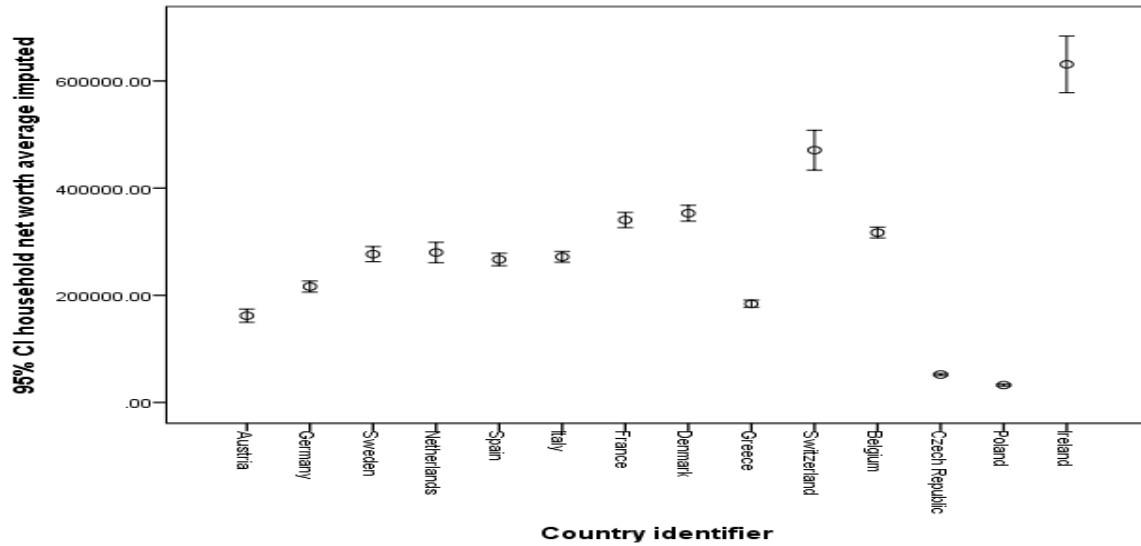
	household net worth average imputed	household income
N	33851	33851
Mean	256977.9025	29221.0135
Std. Error of Mean	2082.95480	229.21091
Median	155404.4063	19524.00
Std. Deviation	383235.3819	42171.69384
Variance	1.469 E+11	1778451762
Range	13615974.75	3185624.50
Minimum	-1815181.75	0.00
Maximum	11800793.00	3185624.50
Percentiles		
25	40803.0039	9194.7871
50	155404.4063	19524.00
75	334134.1563	36092.4219
99	1611439.375	171661.6406

Με βάση τα παραπάνω περιγραφικά στοιχεία παρατηρούμε ότι η ελάχιστη μέση καθαρή αξία τεκμαρτών ενός νοικοκυριού είναι ίση με -1815181.75 και το ελάχιστο εισόδημα ενός νοικοκυριού είναι ίσο με 0. Αντίστοιχα, η μέγιστη μέση καθαρή αξία τεκμαρτών είναι ίση με 11800793.00 και το μέγιστο εισόδημα είναι ίσο με 3185624.50.

Ακόμα, το 75% του πληθυσμού του δείγματος έχει μέση καθαρή αξία τεκμαρτών μέχρι 334134.1563 και εισόδημα νοικοκυριού μέχρι 36092.4219 ενώ το ανώτερο 1% του πληθυσμού έχουν εισόδημα πάνω από 171661.6406.

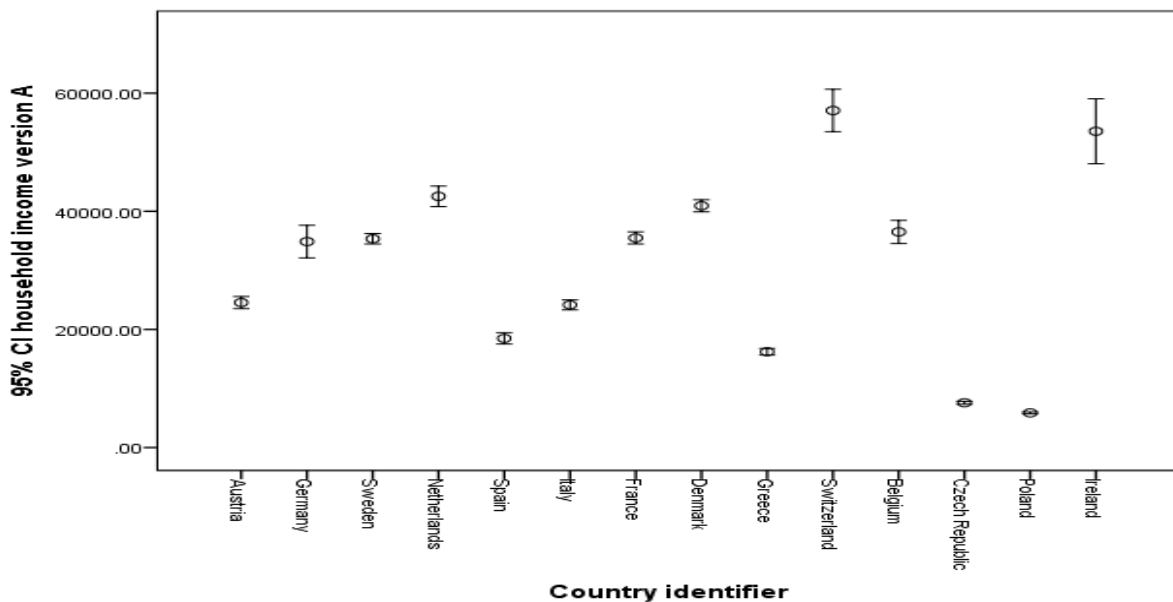
Στη συνέχεια εξετάζουμε τα στοιχεία για τη μέση καθαρή αξία τεκμαρτών και για το εισόδημα νοικοκυριού συγκριτικά για τις 14 χώρες του δείγματος μέσω θηκογράμματος. Όπως φαίνεται στο παρακάτω θηκόγραμμα η Τσέχικη Δημοκρατία και η Πολωνία έχουν εμφανώς χαμηλότερη μέση αξία τεκμαρτών σε σύγκριση με τις υπόλοιπες χώρες.

Διάγραμμα 3.4 Θηκόγραμμα μεταβλητής *hhnetw* συγκριτικά με τη μεταβλητή *country*



Κάτι αντίστοιχο βλέπουμε και στο θηκόγραμμα για το εισόδημα των νοικοκυριών. Παρατηρούμε ότι η Τσέχικη Δημοκρατία και η Πολωνία έχουν χαμηλότερο εισόδημα σε σύγκριση με τις υπόλοιπες 12 χώρες. Ακολουθούν η Ελλάδα, η Ισπανία, η Ιταλία και η Αυστρία που έχουν ελάχιστα καλύτερο εισόδημα. Μέτριο εισόδημα νοικοκυριού φαίνεται ότι έχουν η Γερμανία, η Σουηδία, η Ολλανδία, η Γαλλία, η Δανία και το Βέλγιο και τα υψηλότερα εισοδήματα φαίνεται ότι έχουν η Ελβετία και η Ιρλανδία.

Διάγραμμα 3.5 Θηκόγραμμα μεταβλητής *hhinc* συγκριτικά με τη μεταβλητή *country*



4. ΚΟΙΝΩΝΙΚΟΟΙΚΟΝΟΜΙΚΕΣ ΑΝΙΣΟΤΗΤΕΣ ΚΑΙ ΥΓΕΙΑ

4.1 ΔΕΙΚΤΗΣ ΣΥΓΚΕΝΤΡΩΣΗΣ

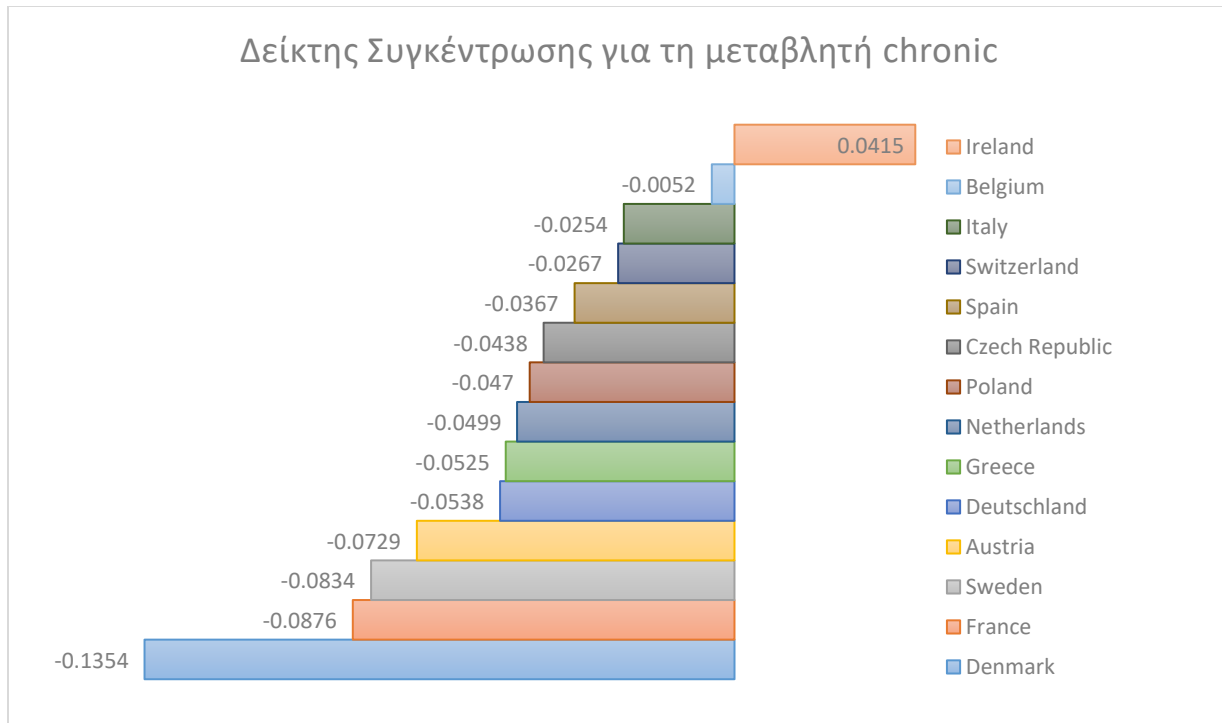
Παρακάτω παρουσιάζουμε τη μέτρηση των ανισοτήτων στην υγεία ως δείκτη συγκέντρωσης (C) όπως τον ορίζει ο Kakwani (2), για τέσσερις μεταβλητές: χρόνιες ασθένειες πάνω από 2 (chronic), περίπτωση κατάθλιψης (eurodcat), αυτοεκτίμηση υγείας (sphus) και περιορισμός στις δραστηριότητες (gali).

Αρχικά δίνεται ο δείκτης συγκέντρωσης καθώς και τα εύρη πιθανών τιμών του, ταξινομώντας τους ερωτώμενους ξεκινώντας με τη λιγότερο μειονεκτική θέση όσον αφορά το εισόδημα των νοικοκυριών (μεταβλητή hhinc).

Πίνακας 4.1 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhinc

chronic	Lower Bound	C	Upper Bound
Ireland	-0.5904	0.0415	0.59241
Belgium	-0.5508	-0.0052	0.55142
Italy	-0.4966	-0.0254	0.49727
Switzerland	-0.7061	-0.0267	0.7075
Spain	-0.5212	-0.0367	0.52209
Czech Republic	-0.5009	-0.0438	0.50169
Poland	-0.4189	-0.047	0.4197
Netherlands	-0.6658	-0.0499	0.66654
Greece	-0.5883	-0.0525	0.58893
Deutschland	-0.59456	-0.0538	0.59534
Austria	-0.5707	-0.0729	0.5724
Sweden	-0.5423	-0.0834	0.54298
France	-0.5821	-0.0876	0.5828
Denmark	-0.5417	-0.1354	0.54245

Διάγραμμα 4.1 Δείκτης Συγκέντρωσης για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhinc

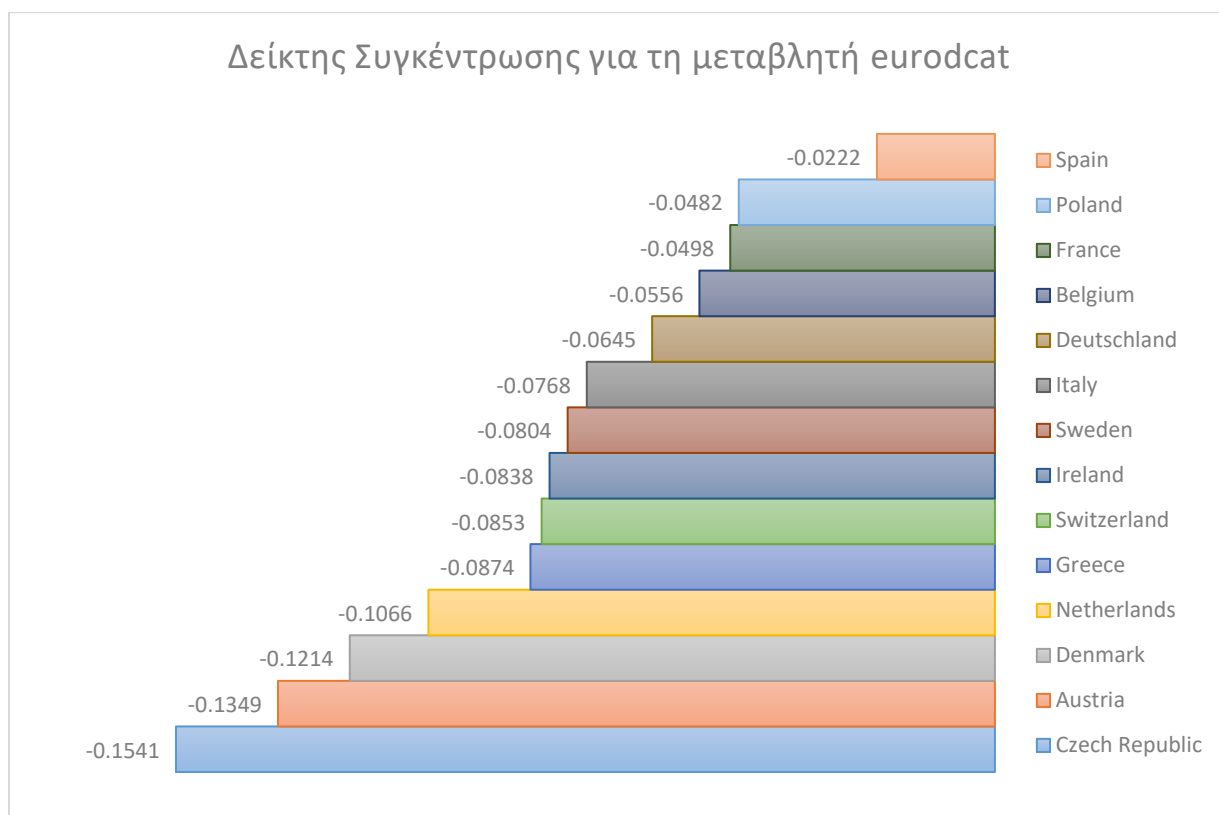


Παρατηρούμε ότι για τη Δανία οι περισσότερες από 2 χρόνιες ασθένειες έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλό εισόδημα νοικοκυριού) ενώ μόνο για την Ιρλανδία η ανισοκατανομή αυτή αντιστρέφεται προς τα υψηλότερα κοινωνικοοικονομικά επίπεδα του. Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Γαλλίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.0876$) έναντι των ερωτηθέντων του Βελγίου ($C=-0.0052$). Γενικά, για όλες τις χώρες πλην της Ιρλανδίας και του Βελγίου οι ερωτηθέντες με περισσότερες από 2 ασθένειες είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.2 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhinc

eurodcat	Lower Bound	C	Upper Bound
Spain	-0.6767	-0.0222	0.67753
Poland	-0.5152	-0.0482	0.51601
France	-0.6814	-0.0498	0.6821
Belgium	-0.7373	-0.0556	0.73798
Deutschland	-0.8218	-0.0645	0.82256
Italy	-0.6636	-0.0768	0.66424
Sweden	-0.835	-0.0804	0.83578
Ireland	-0.817	-0.0838	0.81901
Switzerland	-0.8391	-0.0853	0.8405
Greece	-0.8155	-0.0874	0.81609
Netherlands	-0.8191	-0.1066	0.81987
Denmark	-0.8332	-0.1214	0.834
Austria	-0.7945	-0.1349	0.79625
Czech Republic	-0.776	-0.1541	0.7768

Διάγραμμα 4.2 Δείκτης Συγκέντρωσης για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhinc

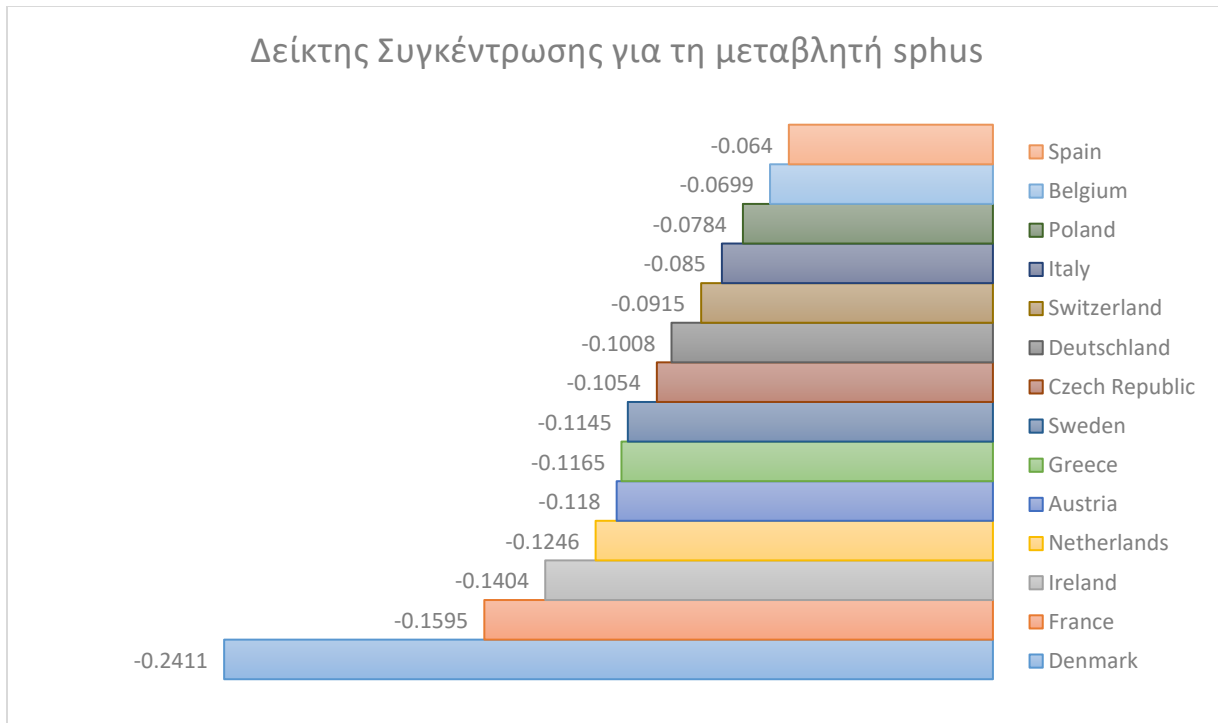


Παρατηρούμε ότι για την Τσέχικη Δημοκρατία οι περιπτώσεις κατάθλιψης των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλό εισόδημα νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Αυστρίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1349$) έναντι των ερωτηθέντων της Ισπανίας ($C=-0.0222$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με εμφάνιση κατάθλιψης είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.3 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhinc

sphus	Lower Bound	C	Upper Bound
Spain	-0.5235	-0.064	0.52435
Belgium	-0.7051	-0.0699	0.7057
Poland	-0.3735	-0.0784	0.37428
Italy	-0.5527	-0.085	0.55339
Switzerland	-0.8204	-0.0915	0.8218
Deutschland	-0.6224	-0.1008	0.62321
Czech Republic	-0.5691	-0.1054	0.56982
Sweden	-0.7059	-0.1145	0.70663
Greece	-0.7354	-0.1165	0.73603
Austria	-0.6624	-0.118	0.66413
Netherlands	-0.7047	-0.1246	0.70546
Ireland	-0.7786	-0.1404	0.78054
France	-0.6301	-0.1595	0.63077
Denmark	-0.7567	-0.2411	0.75746

Διάγραμμα 4.3 Δείκτης Συγκέντρωσης για τη μεταβλητή *sphus* όσον αφορά τη μεταβλητή *hhinc*

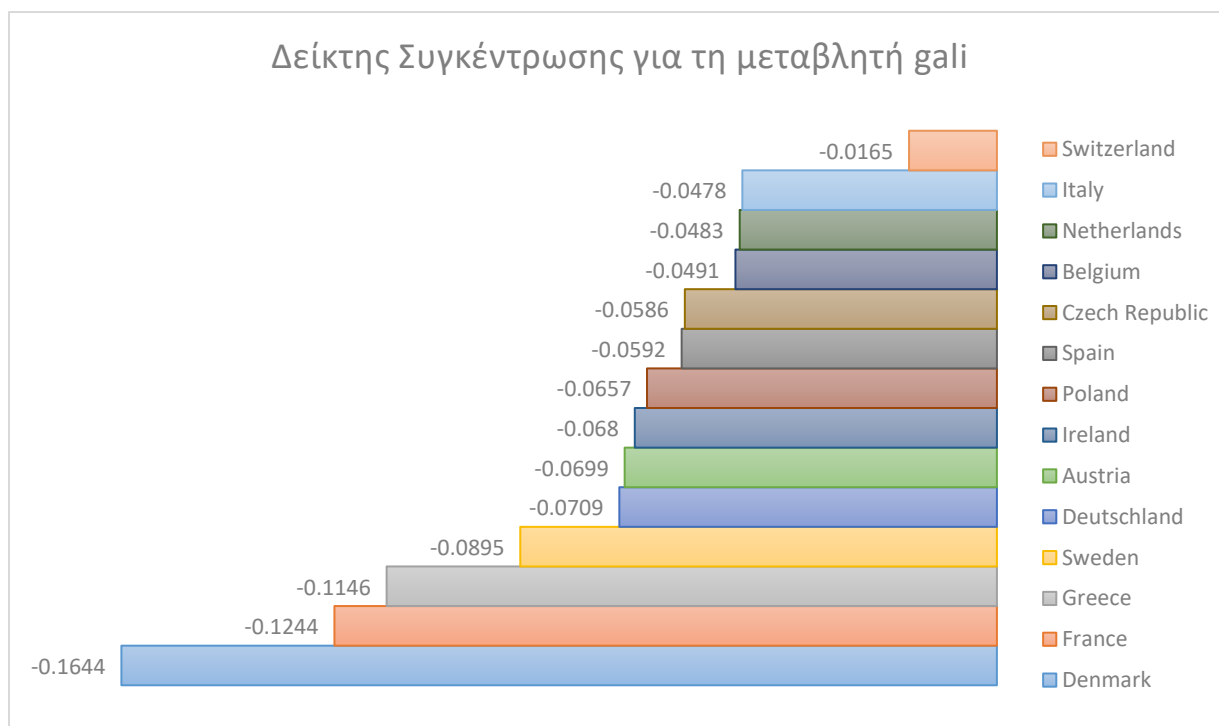


Παρατηρούμε ότι για τη Δανία οι περιπτώσεις χαμηλής αυτοεκτίμησης υγείας των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλό εισόδημα νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Γαλλίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1595$) έναντι των ερωτηθέντων της Ισπανίας ($C=-0.064$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με εμφάνιση χαμηλής αυτοεκτίμησης υγείας είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.4 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή *gali* όσον αφορά τη μεταβλητή *hhinc*

gali	Lower Bound	C	Upper Bound
Switzerland	-0.6825	-0.0165	0.68388
Italy	-0.5748	-0.0478	0.57553
Netherlands	-0.5198	-0.0483	0.52052
Belgium	-0.5939	-0.0491	0.59453
Czech Republic	-0.4227	-0.0586	0.42342
Spain	-0.5714	-0.0592	0.57227
Poland	-0.3831	-0.0657	0.38395
Ireland	-0.6876	-0.068	0.68962
Austria	-0.4776	-0.0699	0.47925
Deutschland	-0.5186	-0.0709	0.51941
Sweden	-0.5723	-0.0895	0.57303
Greece	-0.715	-0.1146	0.71565
France	-0.6202	-0.1244	0.62086
Denmark	-0.6278	-0.1644	0.62854

Διάγραμμα 4.4 Δείκτης Συγκέντρωσης για τη μεταβλητή *gali* όσον αφορά τη μεταβλητή *hhinc*



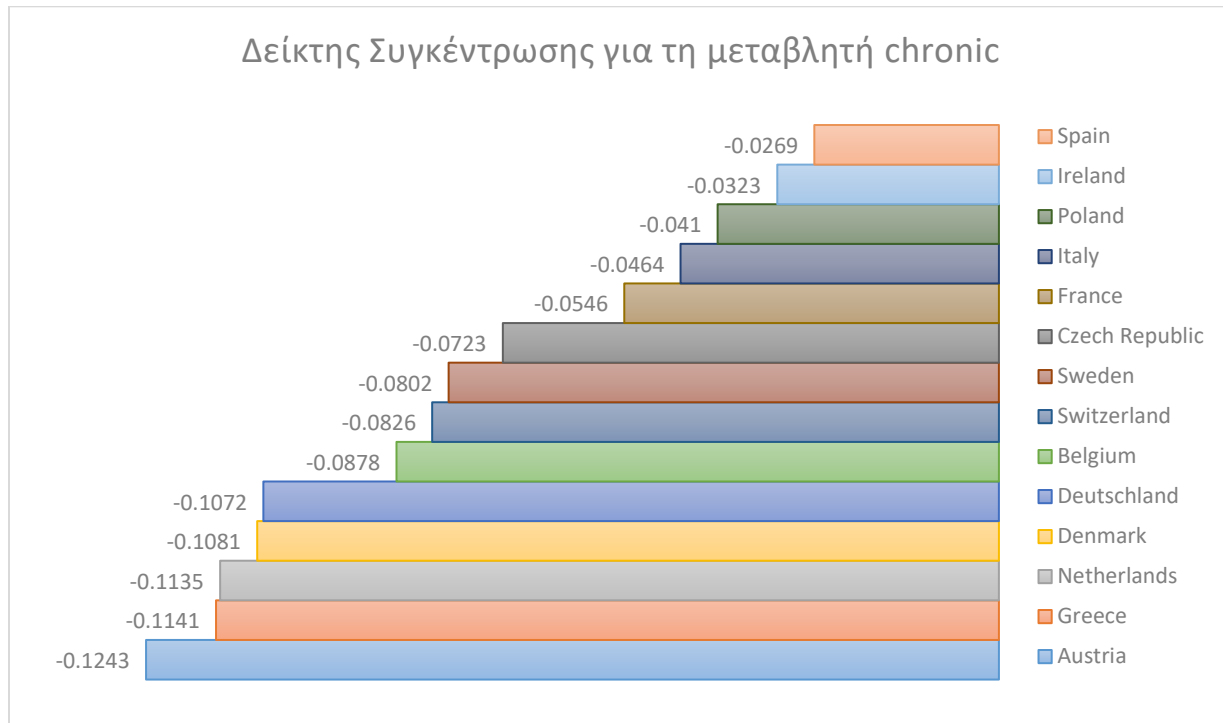
Παρατηρούμε ότι για τη Δανία οι περιπτώσεις περιορισμού των δραστηριοτήτων των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλό εισόδημα νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Γαλλίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1244$) έναντι των ερωτηθέντων της Ελβετίας ($C=-0.0165$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με περιορισμό στις δραστηριότητές τους είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Έπειτα δίνεται ο δείκτης συγκέντρωσης καθώς και τα εύρη πιθανών τιμών του, ταξινομώντας τους ερωτώμενους ξεκινώντας με την πιο μειονεκτική θέση όσον αφορά τη μέση καθαρή αξία τεκμαρτών των νοικοκυριών (μεταβλητή *hhnetw*).

Πίνακας 4.5 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή chronic όσον αφορά τη μεταβλητή hhnetw

chronic	Lower Bound	C	Upper Bound
Spain	-0.5212	-0.0269	0.52209
Ireland	-0.5904	-0.0323	0.59241
Poland	-0.4189	-0.041	0.4197
Italy	-0.4966	-0.0464	0.49727
France	-0.5821	-0.0546	0.5828
Czech Republic	-0.5009	-0.0723	0.50169
Sweden	-0.5423	-0.0802	0.54298
Switzerland	-0.7061	-0.0826	0.7075
Belgium	-0.5508	-0.0878	0.55142
Deutschland	-0.5946	-0.1072	0.59534
Denmark	-0.5417	-0.1081	0.54245
Netherlands	-0.6658	-0.1135	0.66654
Greece	-0.5883	-0.1141	0.58893
Austria	-0.5707	-0.1243	0.5724

Διάγραμμα 4.5 Δείκτης Συγκέντρωσης για τη μεταβλητή *chronic* όσον αφορά τη μεταβλητή *hhnetw*

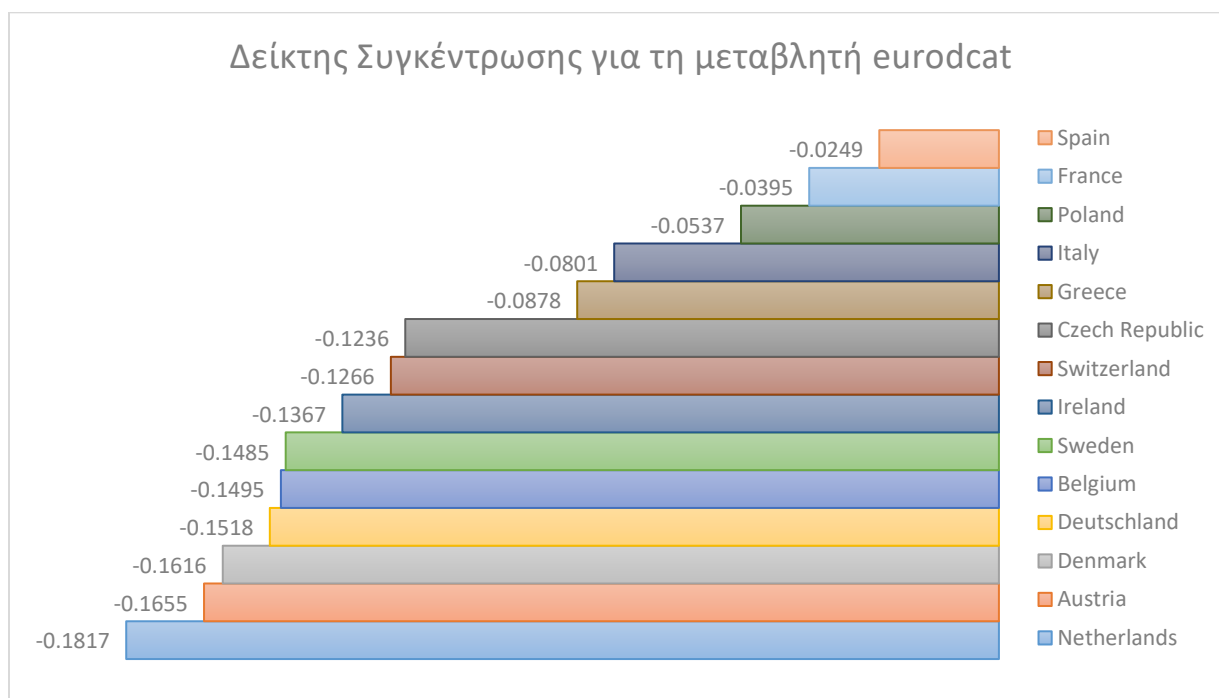


Παρατηρούμε ότι για την Αυστρία οι περισσότερες από 2 χρόνια ασθένειες έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Ελλάδας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1141$) έναντι των ερωτηθέντων της Ισπανίας ($C=-0.0269$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με περισσότερες από 2 ασθένειες είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.6 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhnetw

eurodcat	Lower Bound	C	Upper Bound
Spain	-0.6767	-0.0249	0.67753
France	-0.6814	-0.0395	0.6821
Poland	-0.5152	-0.0537	0.51601
Italy	-0.6636	-0.0801	0.66424
Greece	-0.8155	-0.0878	0.81609
Czech Republic	-0.776	-0.1236	0.7768
Switzerland	-0.8391	-0.1266	0.8405
Ireland	-0.817	-0.1367	0.81901
Sweden	-0.835	-0.1485	0.83578
Belgium	-0.7373	-0.1495	0.73798
Deutschland	-0.8218	-0.1518	0.82256
Denmark	-0.8332	-0.1616	0.834
Austria	-0.7945	-0.1655	0.79625
Netherlands	-0.8191	-0.1817	0.81987

Διάγραμμα 4.6 Δείκτης Συγκέντρωσης για τη μεταβλητή eurodcat όσον αφορά τη μεταβλητή hhnetw

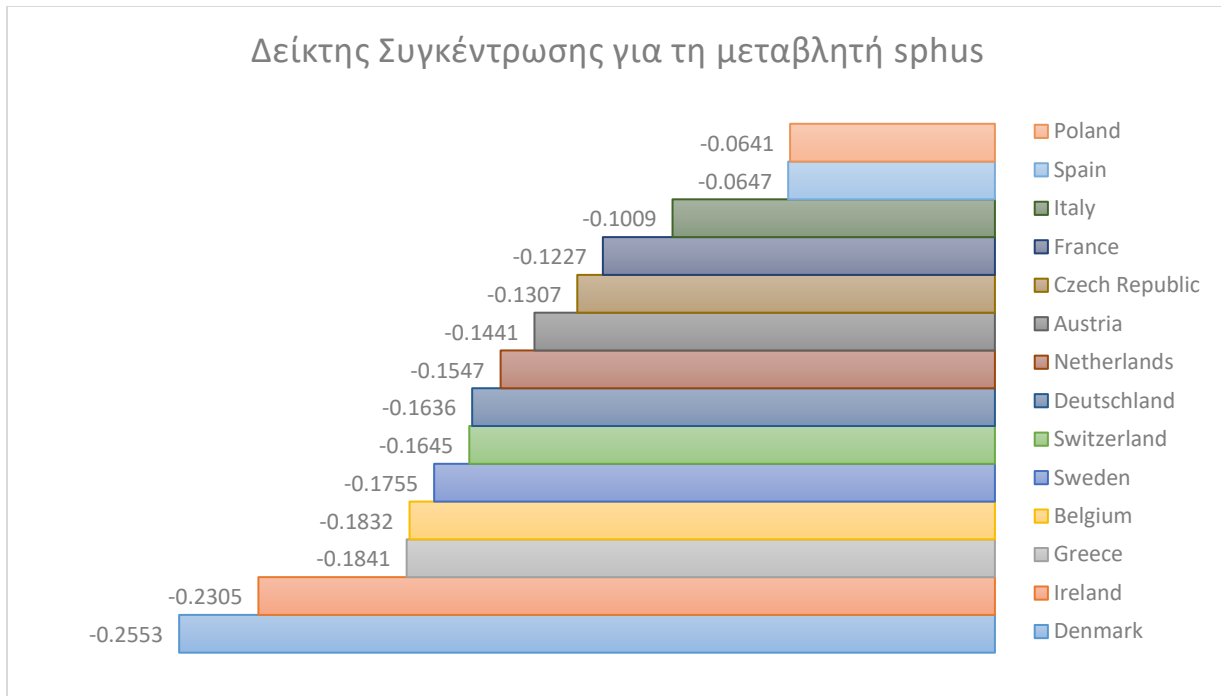


Παρατηρούμε ότι για την Ολλανδία οι περιπτώσεις κατάθλιψης των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Αυστρίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1655$) έναντι των ερωτηθέντων της Ισπανίας ($C=-0.0249$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με εμφάνιση κατάθλιψης είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.7 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή sphus όσον αφορά τη μεταβλητή hhnetw

sphus	Lower Bound	C	Upper Bound
Poland	-0.3735	-0.0641	0.37428
Spain	-0.5235	-0.0647	0.52435
Italy	-0.5527	-0.1009	0.55339
France	-0.6301	-0.1227	0.63077
Czech Republic	-0.5691	-0.1307	0.56982
Austria	-0.6624	-0.1441	0.66413
Netherlands	-0.7047	-0.1547	0.70546
Deutschland	-0.6224	-0.1636	0.62321
Switzerland	-0.8204	-0.1645	0.8218
Sweden	-0.7059	-0.1755	0.70663
Belgium	-0.7051	-0.1832	0.7057
Greece	-0.7354	-0.1841	0.73603
Ireland	-0.7786	-0.2305	0.78054
Denmark	-0.7567	-0.2553	0.75746

Διάγραμμα 4.7 Δείκτης Συγκέντρωσης για τη μεταβλητή *sphus* όσον αφορά τη μεταβλητή *hhnetw*

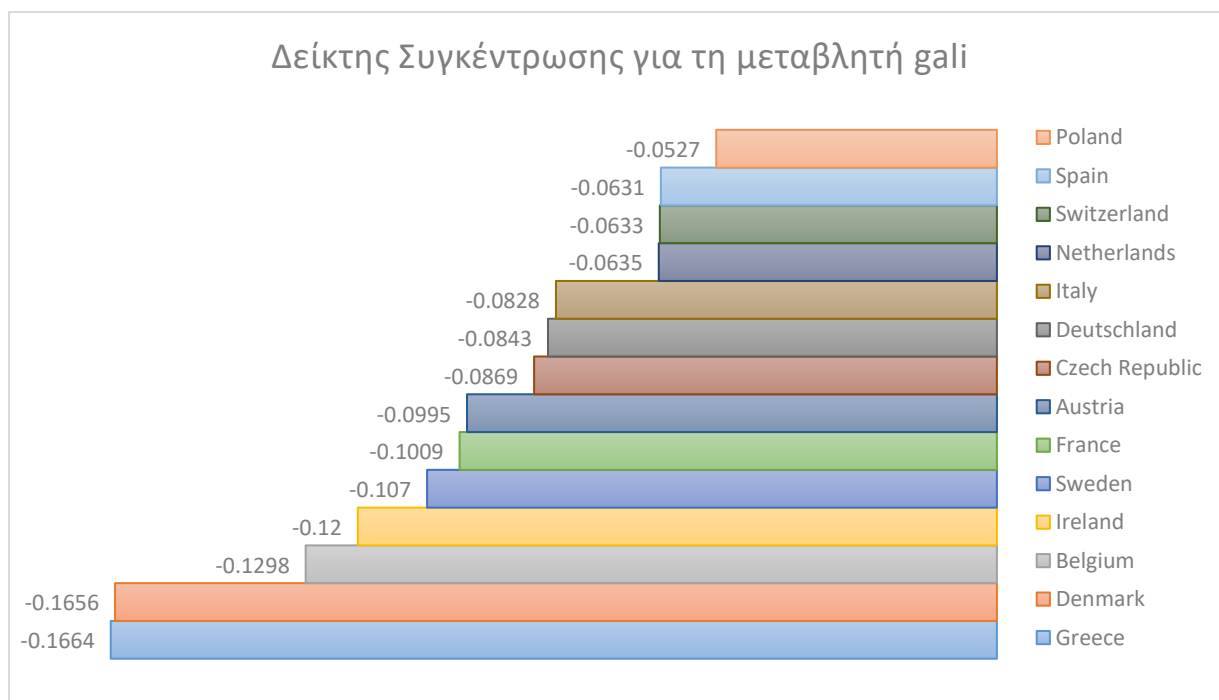


Παρατηρούμε ότι για τη Δανία οι περιπτώσεις χαμηλής αυτοεκτίμησης υγείας των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Ιρλανδίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.2305$) έναντι των ερωτηθέντων της Πολωνίας ($C=-0.0641$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με εμφάνιση χαμηλής αυτοεκτίμησης υγείας είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

Πίνακας 4.8 Δείκτης Συγκέντρωσης και εύρη τιμών του δείκτη για τη μεταβλητή *gali* όσον αφορά τη μεταβλητή *hhnetw*

gali	Lower Bound	C	Upper Bound
Poland	-0.3831	-0.0527	0.38395
Spain	-0.5714	-0.0631	0.57227
Switzerland	-0.6825	-0.0633	0.68388
Netherlands	-0.5198	-0.0635	0.52052
Italy	-0.5748	-0.0828	0.57553
Deutschland	-0.5186	-0.0843	0.51941
Czech Republic	-0.4227	-0.0869	0.42342
Austria	-0.4776	-0.0995	0.47925
France	-0.6202	-0.1009	0.62086
Sweden	-0.5723	-0.107	0.57303
Ireland	-0.6876	-0.12	0.68962
Belgium	-0.5939	-0.1298	0.59453
Denmark	-0.6278	-0.1656	0.62854
Greece	-0.715	-0.1664	0.71565

Διάγραμμα 4.8 Δείκτης Συγκέντρωσης για τη μεταβλητή *gali* όσον αφορά τη μεταβλητή *hhnetw*



Τέλος, παρατηρούμε ότι για την Ελλάδα οι περιπτώσεις περιορισμού των δραστηριοτήτων των ερωτηθέντων έχουν συγκεντρωθεί στο χαμηλότερο κοινωνικοοικονομικό επίπεδο (στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού). Επίσης υπάρχει ανισοκατανομή μεταξύ των ερωτηθέντων της Δανίας με υψηλό κοινωνικοοικονομικό επίπεδο ($C=-0.1656$) έναντι των ερωτηθέντων της Πολωνίας ($C=-0.0527$). Γενικά, για όλες τις χώρες οι ερωτηθέντες με περιορισμό στις δραστηριότητές τους είναι συγκεντρωμένοι στο χαμηλότερο κοινωνικοοικονομικό επίπεδο των πληθυσμών.

4.2 ΛΟΓΙΣΤΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

- Διερεύνηση της επίδρασης κοινωνικοοικονομικών παραγόντων στη μεταβλητή υγείας chronic

Στο παρακάτω μοντέλο διερευνάται πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και η μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_) στη μεταβλητή υγείας που δηλώνει την ύπαρξη ή μη περισσότερων από δύο χρόνιων ασθενειών (chronic). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία. Περισσότεροι πίνακες αποτελεσμάτων δίνονται στο παράρτημα.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.204. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 20.4% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (chronic). Το ποσοστό αυτό είναι ιδιαίτερα ικανοποιητικό. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.6, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 68.8%. Τέλος, οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.599	.080	56.286	1	.000	1.821
yedu_13(1)	-.492	.129	14.538	1	.000	.612
age_group(1)	1.450	.081	320.121	1	.000	4.264
hhnetw_(1)	-.260	.080	10.494	1	.001	.771
Constant	-1.231	.093	176.438	1	.000	.292

a. Variable(s) entered on step 1: gender, yedu_13, age_group, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 82.1% μεγαλύτερες πιθανότητες να εμφανίσουν περισσότερες από δύο χρόνιες ασθένειες σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.821$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες ύπαρξης περισσότερων από δύο χρόνιων ασθενειών κατά 38.8% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.612$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ εμφανίζουν περίπου 4 φορές πιο μεγάλη πιθανότητα ύπαρξης περισσότερων από δύο χρόνιων ασθενειών σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=4.264$.
- Οι ερωτώμενοι που δηλώνουν μέση καθαρή αξία τεκμαρτών του νοικοκυριού μεγαλύτερη της διαμέσου ($hhnetw>123732.6953$) έχουν κατά 22.9% λιγότερες πιθανότητες να εμφανίσουν περισσότερες από δύο χρόνιες ασθένειες σε σύγκριση με τους ερωτώμενους που δηλώνουν μέση καθαρή αξία τεκμαρτών μικρότερη των 123732.6953, με $\exp(b)=0.771$.

Έπειτα παρατίθεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και το εισόδημα του νοικοκυριού (hhinc_) στην ίδια μεταβλητή υγείας (chronic). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.202. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 20.2% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (chronic). Το ποσοστό αυτό είναι ιδιαίτερα ικανοποιητικό. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.229, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 68.8%. Τέλος, οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value}<0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.595	.080	55.215	1	.000	1.814
age_group(1)	1.479	.080	339.561	1	.000	4.390
yedu_13(1)	-.517	.128	16.223	1	.000	.596
hhinc_(1)	-.173	.079	4.728	1	.030	.841
Constant	-1.284	.092	193.784	1	.000	.277

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 81.4% μεγαλύτερες πιθανότητες να εμφανίσουν περισσότερες από δύο χρόνιες ασθένειες σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.814$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες ύπαρξης περισσότερων από δύο χρόνιων ασθενειών κατά 40.4% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.596$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ εμφανίζουν περίπου 4 φορές πιο μεγάλη πιθανότητα ύπαρξης περισσότερων από δύο χρόνιων ασθενειών σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=4.390$.
- Οι ερωτώμενοι που δηλώνουν εισόδημα του νοικοκυριού μεγαλύτερο της διαμέσου ($hhinc>12080.0000$) έχουν κατά 15.9% λιγότερες πιθανότητες να εμφανίσουν περισσότερες από δύο χρόνιες ασθένειες σε σύγκριση με τους ερωτώμενους που δηλώνουν εισόδημα μικρότερο των 12080.0000, με $\exp(b)=0.841$.

➤ Διερεύνηση της επίδρασης κοινωνικοοικονομικών παραγόντων στη μεταβλητή υγείας eurodcat

Στη συνέχεια δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και η μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_) στη μεταβλητή υγείας που δηλώνει την ύπαρξη ή μη κατάθλιψης (eurodcat). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.087. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 8.7% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (eurodcat). Το ποσοστό αυτό δεν είναι ιδιαίτερα ικανοποιητικό. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.920, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του

μοντέλου που αγγίζει το 81.6%. Τέλος, εκτός από τη μεταβλητή που δηλώνει τα χρόνια εκπαίδευσης, οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.947	.103	83.807	1	.000	2.578
age_group(1)	.680	.101	45.695	1	.000	1.973
yedu_13(1)	-.149	.163	.831	1	.362	.862
hhnetw_(1)	-.289	.099	8.527	1	.003	.749
Constant	-2.273	.123	340.609	1	.000	.103

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν περίπου 2.5 φορές πιο μεγάλη πιθανότητα εμφάνισης κατάθλιψης σε σύγκριση με τους άνδρες, αφού $\exp(b)=2.578$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες εμφάνισης κατάθλιψης κατά 13.8% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)=0.862$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν κατά 97.3% μεγαλύτερες πιθανότητες εμφάνισης κατάθλιψης σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=1.973$.
- Οι ερωτώμενοι που δηλώνουν μέση καθαρή αξία τεκμαρτών του νοικοκυριού μεγαλύτερη της διαμέσου ($hhnetw>123732.6953$) έχουν κατά 25.1% λιγότερες πιθανότητες να εμφανίσουν κατάθλιψη σε σύγκριση με τους ερωτώμενους που δηλώνουν μέση καθαρή αξία τεκμαρτών μικρότερη των 123732.6953, με $\exp(b)=0.749$.

Ακόμα δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και το εισόδημα του νοικοκυριού (hhinc_) στην ίδια μεταβλητή υγείας (eurodcat). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.087. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 8.7% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (eurodcat). Το ποσοστό αυτό δεν είναι ιδιαίτερα ικανοποιητικό. Ακόμα, το αποτέλεσμα του test

Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.639, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 81.6%. Τέλος, εκτός από τη μεταβλητή που δηλώνει τα χρόνια εκπαίδευσης, οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.935	.104	81.310	1	.000	2.548
age_group(1)	.704	.099	50.141	1	.000	2.022
yedu_13(1)	-.163	.163	1.001	1	.317	.850
hhinc_(1)	-.276	.098	7.986	1	.005	.759
Constant	-2.285	.122	349.833	1	.000	.102

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν περίπου 2.5 φορές πιο μεγάλη πιθανότητα ύπαρξης κατάθλιψης σε σύγκριση με τους άνδρες, αφού $\exp(b)=2.548$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες ύπαρξης κατάθλιψης κατά 15% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)=0.850$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν περίπου 2 φορές πιο μεγάλη πιθανότητα ύπαρξης κατάθλιψης σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=2.022$.
- Οι ερωτώμενοι που δηλώνουν εισόδημα του νοικοκυριού μεγαλύτερο της διαμέσου ($hhinc>12080.0000$) έχουν κατά 24.1% λιγότερες πιθανότητες ύπαρξης κατάθλιψης σε σύγκριση με τους ερωτώμενους που δηλώνουν εισόδημα μικρότερο των 12080.0000, με $\exp(b)=0.759$.

➤ Διερεύνηση της επίδρασης κοινωνικοοικονομικών παραγόντων στη μεταβλητή υγείας sphus

Παρακάτω δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και η μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_) στη μεταβλητή υγείας που υπολογίζει την αυτοαναφερόμενη υγεία (sphus). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.155. Αυτό σημαίνει ότι το μοντέλο εξηγεί προσεγγιστικά το 15.5% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (sphus). Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.907, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 73.6%. Τέλος, και οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.324	.087	14.042	1	.000	1.383
age_group(1)	1.287	.091	202.013	1	.000	3.622
yedu_13(1)	-.478	.158	9.117	1	.003	.620
hhnetw_(1)	-.467	.088	28.311	1	.000	.627
Constant	-1.655	.104	254.583	1	.000	.191

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 38.3% περισσότερες πιθανότητες να υπολογίσουν την αυτοαναφερόμενη υγεία τους φτωχή σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.383$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες να εκτιμήσουν ως φτωχή την αυτοαναφερόμενη υγεία τους κατά 38% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.620$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν περίπου 3 φορές πιο μεγάλη πιθανότητα να εκτιμήσουν ως φτωχή την αυτοαναφερόμενη υγεία τους σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=3.622$.
- Οι ερωτώμενοι που δηλώνουν μέση καθαρή αξία τεκμαρτών του νοικοκυριού μεγαλύτερη της διαμέσου ($hhnetw>123732.6953$) έχουν κατά 37.3% λιγότερες πιθανότητες να εκτιμήσουν ως φτωχή την αυτοαναφερόμενη υγεία τους σε σύγκριση με τους ερωτώμενους που δηλώνουν μέση καθαρή αξία τεκμαρτών μικρότερη των 123732.6953, με $\exp(b)=0.627$.

Ακόμα δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και το εισόδημα του νοικοκυριού (hhinc_) στην ίδια μεταβλητή υγείας (sphus). Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.153. Αυτό σημαίνει ότι το μοντέλο εξηγεί προσεγγιστικά το 15.3% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (srhus). Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.988, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 73.6%. Τέλος και οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.305	.087	12.303	1	.000	1.356
age_group(1)	1.330	.090	219.286	1	.000	3.782
yedu_13(1)	-.502	.158	10.147	1	.001	.605
hhinc_(1)	-.416	.087	22.927	1	.000	.660
Constant	-1.688	.103	267.925	1	.000	.185

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 35.6% μεγαλύτερη πιθανότητα υπολογισμού της αυτοαναφερόμενης υγείας ως φτωχής σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.356$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες εκτίμησης της υγείας τους ως φτωχής κατά 39.5% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.605$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν περίπου 3 φορές πιο μεγάλη πιθανότητα εκτίμησης της υγείας τους ως φτωχής σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=3.782$.
- Οι ερωτώμενοι που δηλώνουν εισόδημα του νοικοκυριού μεγαλύτερο της διαμέσου ($hhinc > 12080.0000$) έχουν κατά 34% λιγότερες πιθανότητες υπολογισμού της αυτοαναφερόμενης υγείας τους ως φτωχή σε σύγκριση με τους ερωτώμενους που δηλώνουν εισόδημα μικρότερο των 12080.0000, με $\exp(b)=0.660$.

➤ Διερεύνηση της επίδρασης κοινωνικοοικονομικών παραγόντων στη μεταβλητή υγείας gali

Τέλος δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και η μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_) στη μεταβλητή υγείας που υπολογίζει τον περιορισμό των δραστηριοτήτων του ερωτηθέντος (gali). Οι

προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.143. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 14.3% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (gali). Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.584, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 71.5%. Τέλος, και οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.265	.084	9.940	1	.002	1.303
age_group(1)	1.216	.087	194.757	1	.000	3.374
yedu_13(1)	-.498	.151	10.943	1	.001	.608
hhnetw_(1)	-.405	.085	22.592	1	.000	.667
Constant	-1.486	.099	223.299	1	.000	.226

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 30.3% περισσότερες πιθανότητες να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.303$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη κατά 39.2% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.608$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν περίπου 3 φορές πιο μεγάλη πιθανότητα να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=3.374$.
- Οι ερωτώμενοι που δηλώνουν μέση καθαρή αξία τεκμαρτών του νοικοκυριού μεγαλύτερη της διαμέσου ($hhnetw > 123732.6953$) έχουν κατά 33.3% λιγότερες πιθανότητες να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με τους ερωτώμενους που δηλώνουν μέση καθαρή αξία τεκμαρτών μικρότερη των 123732.6953, με $\exp(b)=0.667$.

Ακόμα δίνεται το μοντέλο που διερευνά πώς επιδρούν οι δημογραφικές μεταβλητές για τους Έλληνες ερωτώμενους: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και το εισόδημα του νοικοκυριού (hhinc_) στην ίδια μεταβλητή υγείας (gali). Οι

προαναφερθείσες μεταβλητές είναι κατηγορικές και στο παρακάτω Binary logistic model θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.143. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 14.3% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας (gali). Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.918, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα κάτι το οποίο φαίνεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 72%. Τέλος και οι τέσσερις κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.243	.084	8.305	1	.004	1.275
age_group(1)	1.251	.086	209.514	1	.000	3.493
yedu_13(1)	-.512	.150	11.626	1	.001	.600
hhinc_(1)	-.406	.084	23.147	1	.000	.667
Constant	-1.491	.098	229.278	1	.000	.225

a. Variable(s) entered on step 1: gender, age_group, yedu_13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι

- Οι γυναίκες έχουν κατά 27.5% μεγαλύτερη πιθανότητα να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με τους άνδρες, αφού $\exp(b)=1.275$.
- Για την κατηγορία των ερωτώμενων που έχουν πανεπιστημιακή κατάρτιση (13 χρόνια εκπαίδευσης και πάνω) ισχύει ότι εμφανίζουν λιγότερες πιθανότητες προσδιορισμού της δραστηριότητάς τους ως περιορισμένη κατά 40% από τους ερωτώμενους που έχουν 12 χρόνια εκπαίδευσης και κάτω, με $\exp(b)= 0.600$.
- Οι ερωτώμενοι που ανήκουν στην ηλικιακή κατηγορία 65+ έχουν περίπου 3 φορές πιο μεγάλη πιθανότητα να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με άτομα ηλικίας 50-64, αφού $\exp(b)=3.493$.
- Οι ερωτώμενοι που δηλώνουν εισόδημα του νοικοκυριού μεγαλύτερο της διαμέσου ($hhinc > 12080.0000$) έχουν κατά 33.3% λιγότερες πιθανότητες να προσδιορίσουν τη δραστηριότητά τους ως περιορισμένη σε σύγκριση με τους ερωτώμενους που δηλώνουν εισόδημα μικρότερο των 12080.0000, με $\exp(b)=0.667$.

Τέλος, παρουσιάζονται τα μοντέλα που διερευνούν πώς επιδρούν οι δημογραφικές μεταβλητές για τους ερωτώμενους όλων των χωρών που συμμετείχαν στην έρευνα: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και η μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_) στις τέσσερις μεταβλητές υγείας που αναφέραμε

προηγουμένως, ξεχωριστά. Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στα παρακάτω Binary logistic models θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

- Μεταβλητή ύπαρξης ή μη περισσότερων από δύο χρόνιων ασθενειών (chronic)

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.113. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11.3% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value μικρότερη του 0.001, δηλαδή το μοντέλο δεν προσαρμόζεται ικανοποιητικά στα δεδομένα. Το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 63.6%. Τέλος, εκτός από τέσσερις χώρες (Σουηδία, Γαλλία, Βέλγιο και Ιρλανδία) όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.297	.023	164.249	1	.000	1.346
age_group(1)	.937	.023	1593.958	1	.000	2.553
yedu13(1)	-.276	.028	100.846	1	.000	.759
country			477.886	13	.000	
Austria	-.288	.074	15.025	1	.000	.750
Germany	-.199	.060	10.972	1	.001	.820
Sweden	.008	.058	.020	1	.887	1.008
Netherlands	-.484	.060	64.059	1	.000	.616
Italy	.166	.057	8.372	1	.004	1.180
France	-.094	.058	2.628	1	.105	.910
Denmark	.200	.061	10.680	1	.001	1.221
Greece	-.238	.057	17.559	1	.000	.788
Switzerland	-.662	.073	81.435	1	.000	.516
Belgium	.063	.057	1.233	1	.267	1.065
Czech Republic	.130	.061	4.560	1	.033	1.138
Poland	.428	.062	47.526	1	.000	1.534
Ireland	-.037	.079	.222	1	.638	.963
hhnetw_(1)	-.246	.026	89.733	1	.000	.782
Constant	-.609	.050	148.452	1	.000	.544

a. Variable(s) entered on step 1: gender, age_group, yedu13, country, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από την Αυστρία έχουν λιγότερες πιθανότητες να εμφανίσουν περισσότερες από 2 χρόνιες ασθένειες σε σύγκριση με τους ερωτώμενους από την Ισπανία κατά 25%.

➤ Μεταβλητή ύπαρξης ή μη κατάθλιψης (eurodcat)

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.117. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11.7% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.188, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα πράγμα που επαληθεύεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 76%. Τέλος, εκτός από δύο χώρες (Ιταλία και Γαλλία) όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
country			1217.027	13	.000	
Austria	-.755	.087	75.062	1	.000	.470
Germany	-.778	.071	121.111	1	.000	.459
Sweden	-.868	.071	151.030	1	.000	.420
Netherlands	-.749	.070	115.816	1	.000	.473
Italy	.094	.061	2.346	1	.126	1.098
France	.065	.063	1.065	1	.302	1.067
Denmark	-.736	.073	101.280	1	.000	.479
Greece	-.821	.066	154.032	1	.000	.440
Switzerland	-.865	.087	99.896	1	.000	.421
Belgium	-.186	.063	8.874	1	.003	.830
Czech Republic	-.655	.069	90.213	1	.000	.520
Poland	.551	.065	72.238	1	.000	1.735
Ireland	-.612	.096	40.549	1	.000	.543
gender(1)	.813	.028	833.874	1	.000	2.255
age_group(1)	.285	.027	107.341	1	.000	1.330
yedu13(1)	-.235	.033	49.726	1	.000	.791
hhnetw_(1)	-.340	.031	123.729	1	.000	.712
Constant	-1.146	.056	421.221	1	.000	.318

a. Variable(s) entered on step 1: country, gender, age_group, yedu13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από την Ολλανδία έχουν λιγότερες πιθανότητες ύπαρξης κατάθλιψης σε σύγκριση με τους ερωτώμενους από την Ισπανία κατά 52.7%.

➤ Μεταβλητή υπολογισμού αυτοαναφερόμενης υγείας (sphus)

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.155. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 15.5% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.082, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα πράγμα που επαληθεύεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 68.9%. Τέλος, όλες οι κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
country			1180.980	13	.000	
Spain	-.502	.063	64.122	1	.000	.606
Austria	-1.223	.077	248.974	1	.000	.294
Germany	-.768	.062	153.949	1	.000	.464
Sweden	-1.203	.063	368.523	1	.000	.300
Netherlands	-1.161	.063	337.837	1	.000	.313
Italy	-.553	.060	84.910	1	.000	.575
France	-.743	.061	146.138	1	.000	.476
Denmark	-1.189	.067	312.256	1	.000	.304
Greece	-1.502	.060	620.019	1	.000	.223
Switzerland	-1.774	.084	448.589	1	.000	.170
Belgium	-1.070	.062	299.056	1	.000	.343
Czech Republic	-.735	.059	153.902	1	.000	.480
Ireland	-1.372	.092	224.483	1	.000	.254
gender(1)	.168	.024	47.249	1	.000	1.183
age_group(1)	.763	.025	949.729	1	.000	2.145
yedu13(1)	-.598	.030	392.850	1	.000	.550
hhnetw_(1)	-.453	.027	277.535	1	.000	.636
Constant	.219	.047	21.916	1	.000	1.245

a. Variable(s) entered on step 1: country, gender, age_group, yedu13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από τη Δανία έχουν λιγότερες πιθανότητες υπολογισμού της αυτοαναφερόμενης υγείας τους ως φτωχής σε σύγκριση με τους ερωτώμενους από την Πολωνία κατά 69.6%.

➤ Μεταβλητή ύπαρξης ή μη περιορισμού των δραστηριοτήτων (gali)

Η τιμή του δείκτη Nagelkerke R² είναι ίση με 0.119. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11.9% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value<0.001, δηλαδή το μοντέλο δεν προσαρμόζεται ικανοποιητικά στα δεδομένα. Το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 63.3%. Τέλος, εκτός από την Τσεχία, όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
country			846.553	13	.000	
Spain	-.703	.062	127.355	1	.000	.495
Austria	-.393	.074	27.893	1	.000	.675
Germany	-.354	.061	34.098	1	.000	.702
Sweden	-.605	.060	101.585	1	.000	.546
Netherlands	-.332	.060	30.264	1	.000	.718
Italy	-.658	.060	121.876	1	.000	.518
France	-.737	.061	146.514	1	.000	.478
Denmark	-.635	.063	100.526	1	.000	.530
Greece	-1.363	.059	528.617	1	.000	.256
Switzerland	-1.014	.074	188.914	1	.000	.363
Belgium	-.593	.060	98.819	1	.000	.553
Czech Republic	-.102	.059	2.998	1	.083	.903
Ireland	-.940	.084	123.831	1	.000	.391
gender(1)	.216	.023	86.157	1	.000	1.241
age_group(1)	.762	.024	1042.058	1	.000	2.142
yedu13(1)	-.372	.028	180.389	1	.000	.690
hhnetw_(1)	-.367	.026	199.481	1	.000	.693
Constant	.111	.046	5.774	1	.016	1.117

a. Variable(s) entered on step 1: country, gender, age_group, yedu13, hhnetw_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από την Ελλάδα έχουν λιγότερες πιθανότητες ύπαρξης περιορισμού των δραστηριοτήτων τους σε σύγκριση με τους ερωτώμενους από την Πολωνία κατά 74.4%.

Στη συνέχεια παρατίθενται τα μοντέλα που διερευνούν πώς επιδρούν οι δημογραφικές μεταβλητές για τους ερωτώμενους όλων των χωρών που συμμετείχαν στην έρευνα: φύλο (gender), ηλικιακή ομάδα (age_group), χρόνια εκπαίδευσης (yedu13) καθώς και το εισόδημα του νοικοκυριού (hhinc_) στις ίδιες τέσσερις μεταβλητές υγείας, ξεχωριστά. Οι προαναφερθείσες μεταβλητές είναι κατηγορικές και στα παρακάτω Binary logistic models θέτουμε ως κατηγορία αναφοράς την πρώτη κατηγορία.

- Μεταβλητή ύπαρξης ή μη περισσότερων από δύο χρόνιων ασθενειών (chronic)

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.11. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.019, δηλαδή το μοντέλο δεν προσαρμόζεται ικανοποιητικά στα δεδομένα. Το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 63.5%. Τέλος, εκτός από έξι χώρες (Ισπανία, Αυστρία, Γερμανία, Σουηδία, Γαλλία, και Βέλγιο) όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.296	.023	162.246	1	.000	1.345
age_group(1)	.949	.023	1640.342	1	.000	2.583
yedu13(1)	-.293	.027	113.869	1	.000	.746
country			465.991	13	.000	
Switzerland	-.559	.089	39.121	1	.000	.572
Poland	.565	.081	48.945	1	.000	1.760
Spain	.039	.080	.243	1	.622	1.040
Austria	-.168	.090	3.433	1	.064	.846
Germany	-.075	.079	.911	1	.340	.928
Sweden	.131	.078	2.865	1	.091	1.140
Netherlands	-.365	.079	21.199	1	.000	.694
Italy	.228	.077	8.741	1	.003	1.256
France	-.019	.077	.057	1	.811	.982
Denmark	.308	.079	15.238	1	.000	1.360
Greece	-.151	.077	3.842	1	.050	.860
Belgium	.120	.076	2.456	1	.117	1.127
Czech Republic	.274	.079	11.944	1	.001	1.315
hhinc_(1)	-.147	.027	29.256	1	.000	.863
Constant	-.759	.072	110.938	1	.000	.468

a. Variable(s) entered on step 1: gender, age_group, yedu13, country, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από τη Δανία έχουν μεγαλύτερη πιθανότητα να εμφανίσουν περισσότερες από 2 χρόνιες ασθένειες σε σύγκριση με τους ερωτώμενους από την Ιρλανδία κατά 36%.

➤ **Μεταβλητή ύπαρξης ή μη κατάθλιψης (eurodcat)**

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.116. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11.6% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.143, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα πράγμα που επαληθεύεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 76%. Τέλος, όλες οι κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές $p\text{-value} < 0.05$.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
country			1068.637	13	.000	
Austria	-.620	.087	50.641	1	.000	.538
Germany	-.628	.071	77.946	1	.000	.534
Sweden	-.704	.072	96.187	1	.000	.495
Netherlands	-.597	.071	71.177	1	.000	.551
Italy	.143	.061	5.427	1	.020	1.154
France	.150	.064	5.518	1	.019	1.162
Denmark	-.596	.074	64.438	1	.000	.551
Greece	-.753	.066	131.394	1	.000	.471
Switzerland	-.733	.087	70.203	1	.000	.481
Belgium	-.137	.063	4.730	1	.030	.872
Czech Republic	-.535	.067	63.552	1	.000	.586
Poland	.660	.063	109.316	1	.000	1.936
Ireland	-.590	.096	37.568	1	.000	.554
gender(1)	.804	.028	813.542	1	.000	2.235
age_group(1)	.295	.027	115.760	1	.000	1.343
yedu13(1)	-.244	.033	54.003	1	.000	.783
hhinc_(1)	-.298	.032	88.262	1	.000	.743
Constant	-1.267	.053	563.503	1	.000	.282

a. Variable(s) entered on step 1: country, gender, age_group, yedu13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από την Τσεχία έχουν λιγότερες πιθανότητες ύπαρξης κατάθλιψης σε σύγκριση με τους ερωτώμενους από την Ισπανία κατά 41.4%.

➤ **Μεταβλητή υπολογισμού αυτοαναφερόμενης υγείας (sphus)**

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.153. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 15.3% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value ίση με 0.051, δηλαδή το μοντέλο προσαρμόζεται ικανοποιητικά στα δεδομένα πράγμα που επαληθεύεται και από το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 68.8%. Τέλος, εκτός από τρεις χώρες (Γερμανία, Ιταλία και Τσεχία) όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value < 0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
country			1080.888	13	.000	
Austria	-.532	.077	48.262	1	.000	.587
Germany	-.059	.061	.940	1	.332	.942
Sweden	-.476	.062	58.904	1	.000	.621
Netherlands	-.447	.062	51.230	1	.000	.640
Italy	.017	.058	.090	1	.765	1.017
France	-.118	.060	3.899	1	.048	.889
Denmark	-.494	.067	55.071	1	.000	.610
Greece	-.910	.059	237.547	1	.000	.402
Switzerland	-1.082	.083	170.991	1	.000	.339
Belgium	-.496	.059	69.865	1	.000	.609
Czech Republic	-.081	.059	1.846	1	.174	.922
Poland	.642	.061	109.772	1	.000	1.900
Ireland	-.835	.089	87.091	1	.000	.434
gender(1)	.155	.025	40.066	1	.000	1.168
age_group(1)	.775	.025	983.206	1	.000	2.170
yedu13(1)	-.607	.030	406.957	1	.000	.545
hhinc_(1)	-.418	.028	216.102	1	.000	.658
Constant	-.438	.048	82.329	1	.000	.645

a. Variable(s) entered on step 1: country, gender, age_group, yedu13, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από τη Δανία έχουν λιγότερες πιθανότητες υπολογισμού της αυτοαναφερόμενης υγείας τους ως φτωχής σε σύγκριση με τους ερωτώμενους από την Ισπανία κατά 39%.

➤ Μεταβλητή ύπαρξης ή μη περιορισμού των δραστηριοτήτων (gali)

Η τιμή του δείκτη Nagelkerke R^2 είναι ίση με 0.117. Αυτό σημαίνει ότι το μοντέλο επεξηγεί προσεγγιστικά το 11.7% της μεταβλητότητας της εξαρτημένης μεταβλητής υγείας. Ακόμα, το αποτέλεσμα του test Hosmer-Lemeshow δίνει τιμή p-value<0.001, δηλαδή το μοντέλο δεν προσαρμόζεται ικανοποιητικά στα δεδομένα. Το ποσοστό ορθής ταξινόμησης του μοντέλου που αγγίζει το 63.4%. Τέλος, εκτός από την Ιρλανδία, όλες οι υπόλοιπες κατηγορίες ανεξάρτητων μεταβλητών είναι στατιστικά σημαντικές σε επίπεδο σημαντικότητας 5%, με τιμές p-value<0.05.

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B)
gender(1)	.207	.023	78.812	1	.000	1.230
age_group(1)	.773	.024	1078.502	1	.000	2.167
yedu13(1)	-.382	.028	190.892	1	.000	.683
country			1011.692	13	.000	
Poland	.987	.075	174.302	1	.000	2.682
Spain	.168	.073	5.296	1	.021	1.183
Austria	.627	.084	56.332	1	.000	1.872
Germany	.680	.071	92.261	1	.000	1.974
Sweden	.441	.070	39.504	1	.000	1.554
Netherlands	.702	.070	99.294	1	.000	2.018
Italy	.268	.070	14.627	1	.000	1.307
France	.230	.070	10.666	1	.001	1.258
Denmark	.388	.072	28.964	1	.000	1.475
Greece	-.421	.072	34.437	1	.000	.657
Belgium	.333	.069	23.276	1	.000	1.395
Czech Republic	.895	.073	149.990	1	.000	2.448
Ireland	-.042	.091	.210	1	.647	.959
hhinc_(1)	-.330	.027	145.618	1	.000	.719
Constant	-.889	.065	186.145	1	.000	.411

a. Variable(s) entered on step 1: gender, age_group, yedu13, country, hhinc_.

Από τον τελευταίο πίνακα παρατηρούμε ότι οι ερωτώμενοι από τη Δανία έχουν μεγαλύτερη πιθανότητα ύπαρξης περιορισμού των δραστηριοτήτων τους σε σύγκριση με τους ερωτώμενους από την Ελβετία κατά 47.5%.

5. ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην παρούσα εργασία εξετάσαμε τις κοινωνικοοικονομικές ανισότητες που τυχόν υπάρχουν ανάμεσα σε άτομα ηλικίας 50 ετών και άνω. Εφαρμόσαμε καθιερωμένες μεθόδους στατιστικής ανάλυσης, εστιάζοντας περισσότερο στη Λογιστική Παλινδρόμηση καθώς και μεθόδους εκτίμησης κοινωνικοοικονομικών ανισοτήτων με ιδιαίτερη έμφαση στο Δείκτη Συγκέντρωσης. Το δείγμα των δεδομένων που χρησιμοποιήθηκαν για αυτή την εργασία προέρχεται από το δεύτερο κύμα του Ερευνητικού Προγράμματος SHARE (Survey of Health Ageing and Retirement in Europe) και αποτελείται από 33851 άτομα ηλικίας 50 ετών και άνω τα οποία διακρίνονται σε 14 χώρες.

Σε προηγούμενα κεφάλαια εκτιμήθηκε ο Δείκτης Συγκέντρωσης των δεδομένων μας καθώς επίσης πραγματοποιήθηκε η Λογιστική Παλινδρόμηση για τέσσερις μεταβλητές υγείας οι οποίες υποθέσαμε ότι επηρεάζουν περισσότερο τα άτομα ηλικίας 50 ετών και άνω όσον αφορά το εισόδημα καθώς και τη μέση καθαρή αξία τεκμαρτών των νοικοκυριών (οικονομικές μεταβλητές). Πιο συγκεκριμένα μέσω της Λογιστικής Παλινδρόμησης εξαγάγαμε κάποια συμπεράσματα όσον αφορά το φύλο, την ηλικία, την εκπαίδευση, τις χώρες προέλευσης καθώς και τις δύο οικονομικές μεταβλητές που προαναφέραμε σε σχέση με τις τέσσερις μεταβλητές υγείας ενώ μέσω του Δείκτη Συγκέντρωσης εξαγάγαμε κάποια συμπεράσματα σχετικά με τις οικονομικές μεταβλητές και τις μεταβλητές υγείας.

Από τις ανωτέρω αναλύσεις προκύπτουν τα εξής συμπεράσματα:

1. Όσον αφορά τη μεταβλητή ύπαρξης ή μη περισσότερων από δύο χρόνιων ασθενειών (chronic) σε σχέση με
 - i. τη μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_)

Οι ερωτώμενοι από την Αυστρία έχουν λιγότερες πιθανότητες να εμφανίσουν περισσότερες από 2 χρόνιες ασθένειες, οι οποίες έχουν συγκεντρωθεί στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού, σε σύγκριση με τους ερωτώμενους της Ισπανίας, όπου παρουσιάζεται σχεδόν ισοκατανομή όσον αφορά την εμφάνιση περισσότερων από 2 χρόνιων ασθενειών σε σχέση με την κοινωνικοοικονομική κατάσταση των ερωτηθέντων.

ii. το εισόδημα του νοικοκυριού (hhinc_)

Οι ερωτώμενοι από τη Δανία έχουν μεγαλύτερες πιθανότητες να εμφανίσουν περισσότερες από 2 χρόνιες ασθένειες, οι οποίες έχουν συγκεντρωθεί στους έχοντες χαμηλό εισόδημα νοικοκυριού, σε σύγκριση με τους ερωτώμενους της Ιρλανδίας, όπου έχουν συγκεντρωθεί προς τους έχοντες υψηλό εισόδημα νοικοκυριού όσον αφορά την εμφάνιση περισσότερων από 2 χρόνιων ασθενειών σε σχέση με το κοινωνικοοικονομικό στάτους των ερωτηθέντων.

2. Όσον αφορά τη μεταβλητή ύπαρξης ή μη κατάθλιψης (eurodcat) σε σχέση με

i. τη μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_)

Οι ερωτώμενοι από την Ολλανδία έχουν λιγότερες πιθανότητες ύπαρξης κατάθλιψης, η οποία έχει συγκεντρωθεί στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Ισπανία, όπου παρουσιάζεται σχεδόν ισοκατανομή όσον αφορά την ύπαρξη κατάθλιψης σε σχέση με το κοινωνικοοικονομικό στάτους των ερωτηθέντων.

ii. το εισόδημα του νοικοκυριού (hhinc_)

Οι ερωτώμενοι από την Τσέχικη Δημοκρατία έχουν λιγότερες πιθανότητες ύπαρξης κατάθλιψης, η οποία έχει συγκεντρωθεί στους έχοντες χαμηλό εισόδημα νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Ισπανία, όπου παρουσιάζεται σχεδόν ισοκατανομή όσον αφορά την ύπαρξη κατάθλιψης σε σχέση με το κοινωνικοοικονομικό στάτους των ερωτηθέντων.

3. Όσον αφορά τη μεταβλητή υπολογισμού αυτοαναφερόμενης υγείας (srhus) σε σχέση με

i. τη μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_)

Οι ερωτώμενοι από τη Δανία έχουν λιγότερες πιθανότητες εκτίμησης ως φτωχής της υγείας τους, γεγονός το οποίο έχει συγκεντρωθεί στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Πολωνία, όπου παρουσιάζεται λιγότερη συγκέντρωση αυτοαναφερόμενης ως φτωχής υγείας στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών νοικοκυριού.

ii. το εισόδημα του νοικοκυριού (hhinc_)

Οι ερωτώμενοι από τη Δανία έχουν λιγότερες πιθανότητες εκτίμησης ως φτωχής της υγείας τους, γεγονός το οποίο έχει συγκεντρωθεί στους έχοντες χαμηλό εισόδημα νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Ισπανία, όπου παρουσιάζεται λιγότερη συγκέντρωση αυτοαναφερόμενης ως φτωχής υγείας στους έχοντες χαμηλό εισόδημα νοικοκυριού.

4. Όσον αφορά τη μεταβλητή ύπαρξης ή μη περιορισμού των δραστηριοτήτων (gali) σε σχέση με

i. τη μέση καθαρή αξία τεκμαρτών του νοικοκυριού (hhnetw_)

Οι ερωτώμενοι από την Ελλάδα έχουν λιγότερες πιθανότητες ύπαρξης περιορισμού των δραστηριοτήτων τους, γεγονός το οποίο έχει συγκεντρωθεί στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών του νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Πολωνία, όπου παρουσιάζεται λιγότερη συγκέντρωση ύπαρξης περιορισμού των δραστηριοτήτων στους έχοντες χαμηλή μέση καθαρή αξία τεκμαρτών νοικοκυριού.

ii. το εισόδημα του νοικοκυριού (hhinc_)

Οι ερωτώμενοι από τη Δανία έχουν μεγαλύτερη πιθανότητα ύπαρξης περιορισμού των δραστηριοτήτων τους, γεγονός το οποίο έχει συγκεντρωθεί στους έχοντες χαμηλό εισόδημα νοικοκυριού, σε σύγκριση με τους ερωτώμενους από την Ελβετία, όπου παρουσιάζεται σχεδόν ισοκατανομή όσον αφορά την ύπαρξη περιορισμού των δραστηριοτήτων σε σχέση με το κοινωνικοοικονομικό στάτους των ερωτηθέντων.

ΠΑΡΑΡΤΗΜΑ

Πίνακας 0.1 Κωδικοποίηση του “CC” για τη μεταβλητή mergeid

AT	Austria (Αυστρία)	DK	Denmark (Δανία)
DE	Germany (Γερμανία)	GR	Greece (Ελλάδα)
SE	Sweden (Σουηδία)	Cg, Cf, Ci	Switzerland (Ελβετία)
NL	Netherlands (Ολλανδία)	Bf, Bn	Belgium (Βέλγιο)
ES	Spain (Ισπανία)	CZ	Czech Republic (Τσεχία)
IT	Italy (Ιταλία)	PL	Poland (Πολωνία)
FR	France (Γαλλία)	IE	Ireland (Ιρλανδία)

Πίνακας 0.2 Κωδικοποίηση χωρών για τη μεταβλητή country

11	Austria (Αυστρία)	18	Denmark (Δανία)
12	Germany (Γερμανία)	19	Greece (Ελλάδα)
13	Sweden (Σουηδία)	20	Switzerland (Ελβετία)
14	Netherlands (Ολλανδία)	23	Belgium (Βέλγιο)
15	Spain (Ισπανία)	28	Czech Republic (Τσεχία)
16	Italy (Ιταλία)	29	Poland (Πολωνία)
17	France (Γαλλία)	30	Ireland (Ιρλανδία)

Πίνακας 0.3 Κωδικοποίηση μεταβλητής chronic2w2

0	λιγότερο από 2 ασθένειες	-2	Αρνήθηκαν να απαντήσουν
1	2 ή περισσότερες ασθένειες	-1	Δε γνωρίζω

Πίνακας 0.4 Στοιχεία της μεταβλητής υγείας eurodcat

euro1	κατάθλιψη	euro7	ευερεθιστικότητα
euro2	απαισιοδοξία	euro8	όρεξη
euro3	τάση αυτοκτονίας	euro9	κόπωση
euro4	ενοχή	euro10	συγκέντρωση (ανάγνωση και ψυχαγωγία)
euro5	ύπνος	euro11	απόλαυση
euro6	ενδιαφέρον	euro12	δάκρυα

Πίνακας 0.5 Κωδικοποίηση μεταβλητής *gender*

1	Άντρας	-2	Αρνήθηκαν να απαντήσουν
2	Γυναίκα	-1	Δε γνωρίζω

Πίνακας 0.6 Κωδικοποίηση μεταβλητής *ageclass*

1	50-54
2	55-59
3	60-64
4	65-69
5	70-74
6	75-79
7	80-84
8	85-89
9	90-94
10	95-99
11	100-104

Πίνακας 0.7 Κωδικοποίηση της μεταβλητής *sphus*

1	εξαιρετική	4	μέτρια
2	πολύ καλή	5	φτωχή
3	καλή	-99	δε συμπληρώθηκε

Πίνακας 0.8 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *chronic* για τους Έλληνες ερωτούμενους και τη μεταβλητή *hhnetw*

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3829.405 ^a	.151	.204

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.493	7	.600

Classification Table^a

Observed		Predicted		
		2+ chronic diseases (w2 version)		Percentage Correct
		Less than 2 diseases	2+ chronic diseases	
2+ chronic diseases (w2 version)	Less than 2 diseases 2+ chronic diseases	1466 576	427 747	77.4 56.5
Overall Percentage				68.8

a. The cut value is .500

Πίνακας 0.9 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3835.137 ^a	.150	.202

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	9.346	7	.229

Classification Table^a

Observed		Predicted		
		2+ chronic diseases (w2 version)		Percentage Correct
		Less than 2 diseases	2+ chronic diseases	
2+ chronic diseases (w2 version)	Less than 2 diseases	1334	559	70.5
	2+ chronic diseases	444	879	66.4
Overall Percentage				68.8

a. The cut value is .500

Πίνακας 0.10 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους Έλληνες ερωτώμενους και τη μεταβλητή hhnetw

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2819.539 ^a	.054	.087

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	3.220	8	.920

Classification Table^a

Observed		Predicted		
		EURO-D caseness		Percentage Correct
		No	Yes	
EURO-D caseness	No	2555	0	100.0
	Yes	577	0	.0
Overall Percentage				81.6

a. The cut value is .500

Πίνακας 0.11 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *eurodcat* για τους Έλληνες ερωτώμενους και τη μεταβλητή *hhinc*

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	2820.081 ^a	.054	.087

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.170	7	.639

Classification Table^a

Observed		Predicted		
		EURO-D caseness		Percentage Correct
		No	Yes	
EURO-D caseness	No	2555	0	100.0
	Yes	577	0	.0
Overall Percentage				81.6

a. The cut value is .500

Πίνακας 0.12 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *sphus* για τους Έλληνες ερωτώμενους και τη μεταβλητή *hhnetw*

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3376.213 ^a	.106	.155

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.748	7	.907

Classification Table^a

Observed		Predicted		
		sphus_new		Percentage Correct
		0	1	
sphus_new	0	2383	0	100.0
	1	856	0	.0
Overall Percentage				73.6

a. The cut value is .500

Πίνακας 0.13 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας sphus για τους Έλληνες ερωτώμενους και τη μεταβλητή hhinc

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3381.637 ^a	.105	.153

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	1.318	7	.988

Classification Table^a

Observed		Predicted		
		sphus_new		Percentage Correct
		0	1	
sphus_new	0	2383	0	100.0
	1	856	0	.0
Overall Percentage				73.6

a. The cut value is .500

Πίνακας 0.14 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *gali* για τους Έλληνες ερωτώμενους και τη μεταβλητή *hhnetw*

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3528.855 ^a	.100	.143

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	5.624	7	.584

Classification Table^a

Observed		Predicted		
		Limitations with activities - gali		Percentage Correct
		Not limited	Limited	
Limitations with activities - gali	Not limited	2317	0	100.0
	Limited	922	0	.0
Overall Percentage				71.5

a. The cut value is .500

Πίνακας 0.15 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *gali* για τους Έλληνες ερωτώμενους και τη μεταβλητή *hhinc*

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	3528.255 ^a	.100	.143

a. Estimation terminated at iteration number 5 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	2.620	7	.918

Classification Table^a

Observed		Predicted		
		Limitations with activities - gali		Percentage Correct
		Not limited	Limited	
Limitations with activities - gali	Not limited	2045	272	88.3
	Limited	636	286	31.0
Overall Percentage				72.0

a. The cut value is .500

Πίνακας 0.16 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας chronic για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw

		Frequency	Parameter coding												
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Country identifier	Spain	2377	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1181	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2575	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2757	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2630	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2926	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	France	2838	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Denmark	2544	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Greece	3216	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Switzerland	1453	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Belgium	3141	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Czech Republic	2663	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Poland	2416	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Ireland	1001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
	Gender	0	16162	.000											
1		17556	1.000												
0		17455	.000												
1		16263	1.000												
0		23737	.000												
yedul3	1	9981	1.000												
	Male	15210	.000												
	Female	18508	1.000												

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	43348.206 ^a	.084	.113

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	29.228	8	.000

Classification Table^a

Observed		Predicted		
		2+ chronic diseases (w2 version)		Percentage Correct
		Less than 2 diseases	2+ chronic diseases	
2+ chronic diseases (w2 version)	Less than 2 diseases 2+ chronic diseases	13421 6925	5351 8021	71.5 53.7
Overall Percentage				63.6

a. The cut value is .500

Πίνακας 0.17 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw

Categorical Variables Codings

	Frequency	Parameter coding												
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Country identifier	Spain	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2536	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2661	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2587	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2889	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	France	2699	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Denmark	2494	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Greece	3132	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Switzerland	1442	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Belgium	3099	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Czech Republic	2603	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Poland	2343	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
Ireland	989	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	
Gender	Male	14844	.000											
	Female	18073	1.000											
hhnetw_	0	15672	.000											
	1	17245	1.000											
yedu13	0	23073	.000											
	1	9844	1.000											
	0	17197	.000											
age_group	1	15720	1.000											

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	34015.203 ^a	.079	.117

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	11.247	8	.188

Classification Table^a

Observed		Predicted		
		EURO-D caseness		Percentage Correct
		No	Yes	
EURO-D caseness	No	24074	751	97.0
	Yes	7138	954	11.8
Overall Percentage				76.0

a. The cut value is .500

Πίνακας 0.18 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *sphus* για τους ερωτώμενους όλων των χωρών και τη μεταβλητή *hhnetw*

		Categorical Variables Codings													
		Frequency	Parameter coding												
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Country identifier	Poland	2426	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Spain	2382	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1182	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2585	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2761	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2638	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2931	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	France	2879	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Denmark	2548	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Greece	3239	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Switzerland	1459	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Belgium	3143	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Czech Republic	2671	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Ireland	1007	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
	Gender	Male	15270	.000											
	Female	18581	1.000												
hhnetw_	0	16219	.000												
	1	17632	1.000												
yedu13	0	23840	.000												
	1	10011	1.000												
age_group	0	17520	.000												
	1	16331	1.000												

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	39994.557 ^a	.113	.155

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	14.004	8	.082

Classification Table^a

Observed		Predicted		
		sphus_new		Percentage Correct
		0	1	
sphus_new	0	19382	2446	88.8
	1	8066	3957	32.9
Overall Percentage				68.9

a. The cut value is .500

Πίνακας 0.19 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhnetw

Categorical Variables Codings

	Frequency	Parameter coding															
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)			
Country identifier	Poland	2417	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Spain	2380	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1181	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2576	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2759	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2632	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2926	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	France	2838	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	Denmark	2544	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Greece	3239	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Switzerland	1452	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Belgium	3142	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Czech Republic	2664	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
Ireland	1002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	
Gender	Male	15223	.000														
	Female	18529	1.000														
hhnetw_	0	16186	.000														
	1	17566	1.000														
yedu13	0	23769	.000														
	1	9983	1.000														
	0	17465	.000														
age_group	0	17465	.000														
	1	16287	1.000														

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	43033.615 ^a	.089	.119

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	28.279	8	.000

Classification Table^a

Observed		Predicted		
		Limitations with activities - gali		Percentage Correct
		Not limited	Limited	
Limitations with activities - gali	Not limited	14404	4770	75.1
	Limited	7623	6955	47.7
Overall Percentage				63.3

a. The cut value is .500

Πίνακας 0.20 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *chronic* για τους ερωτώμενους όλων των χωρών και τη μεταβλητή *hhinc*

		Categorical Variables Codings														
		Frequency	Parameter coding													
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	
Country identifier	Ireland	1001	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Switzerland	1453	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Poland	2416	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Spain	2377	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Austria	1181	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Germany	2575	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	
	Sweden	2757	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	
	Netherlands	2630	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	
	Italy	2926	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	
	France	2838	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	
	Denmark	2544	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	
	Greece	3216	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	
	Belgium	3141	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	
	Czech Republic	2663	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	
	age_group	0	16868	.000												
		1	16850	1.000												
		0	17455	.000												
yedu13	1	16263	1.000													
	0	23737	.000													
Gender	Male	9981	1.000													
	Female	15210	.000													
	Female	18508	1.000													

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	43408.642 ^a	.082	.110

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	18.317	8	.019

Classification Table^a

Observed		Predicted		
		2+ chronic diseases (w2 version)		Percentage Correct
		Less than 2 diseases	2+ chronic diseases	
2+ chronic diseases (w2 version)	Less than 2 diseases	13312	5460	70.9
	2+ chronic diseases	6853	8093	54.1
Overall Percentage				63.5

a. The cut value is .500

Πίνακας 0.21 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας eurodcat για τους ερωτούμενους όλων των χωρών και τη μεταβλητή hhinc

Categorical Variables Codings

	Frequency	Parameter coding												
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Country identifier														
Spain	2270	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Austria	1173	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Germany	2536	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Sweden	2661	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Netherlands	2587	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
Italy	2889	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
France	2699	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
Denmark	2494	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
Greece	3132	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
Switzerland	1442	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
Belgium	3099	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
Czech Republic	2603	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
Poland	2343	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
Ireland	989	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
Gender														
Male	14844	.000												
Female	18073	1.000												
hhinc_														
0	16365	.000												
1	16552	1.000												
yedu13														
0	23073	.000												
1	9844	1.000												
age_group														
0	17197	.000												
1	15720	1.000												

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	34050.701 ^a	.078	.116

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	12.200	8	.143

Classification Table^a

Observed		Predicted		
		EURO-D caseness		Percentage Correct
		No	Yes	
EURO-D caseness	No	24344	481	98.1
	Yes	7417	675	8.3
Overall Percentage				76.0

a. The cut value is .500

Πίνακας 0.22 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας *sphus* για τους ερωτώμενους όλων των χωρών και τη μεταβλητή *hhinc*

		Categorical Variables Codings													
		Frequency	Parameter coding												
			(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
Country identifier	Spain	2382	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1182	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2585	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2761	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2638	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2931	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	France	2879	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Denmark	2548	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	Greece	3239	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Switzerland	1459	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Belgium	3143	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Czech Republic	2671	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Poland	2426	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
	Ireland	1007	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000
Gender	Male	15270	.000												
	Female	18581	1.000												
hhinc_	0	16926	.000												
	1	16925	1.000												
yedul3	0	23840	.000												
	1	10011	1.000												
age_group	0	17520	.000												
	1	16331	1.000												

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	40056.290 ^a	.111	.153

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	15.466	8	.051

Classification Table^a

Observed		Predicted		
		sphus_new		Percentage Correct
		0	1	
sphus_new	0	19156	2672	87.8
	1	7890	4133	34.4
Overall Percentage				68.8

a. The cut value is .500

Πίνακας 0.23 Κατάλογος πινάκων για το μοντέλο λογιστικής παλινδρόμησης της μεταβλητής υγείας gali για τους ερωτώμενους όλων των χωρών και τη μεταβλητή hhinc

Categorical Variables Codings

	Frequency	Parameter coding													
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	
Country identifier	Switzerland	1452	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Poland	2417	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Spain	2380	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Austria	1181	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Germany	2576	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000	.000
	Sweden	2759	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000	.000
	Netherlands	2632	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000	.000
	Italy	2926	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000	.000
	France	2838	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000	.000
	Denmark	2544	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000	.000
	Greece	3239	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000	.000
	Belgium	3142	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000	.000
	Czech Republic	2664	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	.000
Ireland	1002	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	.000	1.000	
hhinc_	0	16899	.000												
	1	16853	1.000												
age_group	0	17465	.000												
	1	16287	1.000												
yedu13	0	23769	.000												
	1	9983	1.000												
Gender	Male	15223	.000												
	Female	18529	1.000												

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	43087.059 ^a	.087	.117

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	27.145	8	.001

Classification Table^a

Observed		Predicted		
		Limitations with activities - gali		Percentage Correct
		Not limited	Limited	
Limitations with activities - gali	Not limited	14257	4917	74.4
	Limited	7450	7128	48.9
Overall Percentage				63.4

a. The cut value is .500

ΒΙΒΛΙΟΓΡΑΦΙΑ

- Borsch-Supan, A., Brugiavini, A., Jorges, H., Mackenbach, J., Siegrist, J., & Weber, G. (2005). SHARE: First results from the survey of Health, Ageing and Retirement in Europe.
- Clarke, P. M., Gerdtham, U.-G., & Connelly, L. B. (2002, December 23). A note on the decomposition of the health concentration index. *Health Economics*, pp. 511-516.
- Doorslaer, E. v., Koolman, X., & Jones, A. M. (2004, June 14). Explaining income-related inequalities in doctor utilisation. *Health Economics* 13, pp. 629-647.
- HERNÁNDEZ-QUEVEDO, C., & MASSERIA, C. (2013). Measuring Income-Related Inequalities in Health in Multi-Country Analysis. *ESTUDIOS DE ECONOMÍA APLICADA*, pp. 445-476.
- Kakwani, N., Wagstaff, A., & Doorslaer, E. v. (1997). Socioeconomic inequalities in health: Measurement, computation, and statistical inference. *Journal of Econometrics* 77, pp. 87-103.
- Laaksonen, M., Rahkonen, O., Karvonen, S., & Lahelma, E. (2005, March 8). Socioeconomic status and smoking: Analyzing inequalities with multiple indicators. *European Journal of Public Health* 15, pp. 262-269.
- Mackenbach, J. P., & Kunst, A. E. (1997). Measuring the magnitude of socio-economic inequalities in health: an overview of available measures illustrated with two examples from Europe. *Social Science and Medicine* 44, pp. 757-771.
- Technical Note, 7. (n.d.). *The Concentration Index*. Quantitative Techniques for Health Equity Analysis.
- Tsimbos, C. (2010). An assessment of socio-economic inequalities in health among elderly in Greece, Italy and Spain. *Public Health* 55, pp. 5-15.
- Wagstaff, A. (2004, October 20). The bounds of the concentration index when the variable of interest is binary, with an application to immunization inequality. *Health Economics* 14, pp. 429-432.
- Wagstaff, A., Doorslaer, E. v., & Watanabe, N. (2003). On decomposing the causes of health sector inequalities with an application to malnutrition inequalities in Vietnam. *Journal of Econometrics* 112, pp. 207-223.
- Wagstaff, A., Paci, P., & Doorslaer, E. v. (1991). On the measurement of inequalities in health. *Social Science and Medicine* 33, pp. 545-557.
- Yiengprugsawan, V., Lim, L. L., Carmichael, G. A., Dear, K. B., & Sleigh, A. C. (2010). Decomposing socioeconomic inequality for binary health outcomes: an improved estimation that does not vary by choice of reference group. *BMC Research Notes*, pp. 57-61.
- Αναστασιάδου, Σ. Δ. (2012). *Στατιστική και μεθοδολογία έρευνας στις κοινωνικές επιστήμες*. Εκδόσεις Κριτική.
- Δουληγέρης, Β., Καππή, Χ., & Μαγδαληνός, Μ. (1997). *Κοινωνικοοικονομικοί προσδιοριστικοί παράγοντες της γονιμότητας στην Ελλάδα*. Εθνικό Κέντρο Κοινωνικών Ερευνών.
- Καλογήρου, Σ., Τραγάκη, Α., Τσίμπος, Κ., & Μουστάκη, Ε. (2011). *Χωρικές ανισότητες εισοδήματος, ανάπτυξης και φτώχειας στην Ελλάδα*. Αθήνα: Πρόγραμμα Επιστημών Μελετών.

- Κατσαμάγκος, Θ. (2007). *Κοινωνικές και οικονομικές διαφορές και ανισότητες μεταξύ των ηλικιωμένων στην Ε.Ε.*
- Κοινωνικές ανισότητες και κοινωνικός αποκλεισμός. (1996). Στο *6ο Επιστημονικό Συνέδριο 27-30 Νοεμβρίου*. Ίδρυμα Σάκη Καράγιωργα.
- Λυμπεράκη, Α., Τήνιος, Π., & Φιλαλήθης, Τ. (2009). *Υγεία, γήρανση και σύνταξη στην Ελλάδα και στην Ευρώπη, Ζωή 50+*. Κριτική.
- Μητούλα, Ρ. (2006). *Βιώσιμη περιφερειακή ανάπτυξη στην Ευρωπαϊκή Ένωση και ανασυγκρότηση του Ελληνικού αστικού περιβάλλοντος*. Εκδόσεις Σταμούλης.
- Παπαευαγγέλου, Γ., & Τσίμπος, Κ. (1983). *Κοινωνικές και υγειονομικές επιπτώσεις από τις δημογραφικές εξελίξεις του Ελληνικού πληθυσμού*. Αθήνα: Υγειονομική Σχολή Αθηνών.
- Πετμετζίδου-Τσουλουβή, Μ. (1992). *Κοινωνικές ανισότητες και κοινωνική πολιτική*. Εκδότης Εξάντας.
- Σαββάκης, Μ. (2013). *Μικροκοινωνιολογία και ποιοτική έρευνα*. Εκδόσεις Κριτική.
- Φίλιας, Β. Ι. (2002). *Κοινωνιολογία της οικονομίας*. Gutenberg - Γιώργος & Κώστας Δαρδανός.
- Χουμανίδης, Λ. Θ. (1978). *Εισαγωγή εις την οικονομικήν κοινωνιολογίαν*. Πειραιάς: Στ. Καραμπερόπουλος.