



ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ
ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ
ΕΠΙΣΤΗΜΩΝ

Διπλωματική Εργασία:
**ΕΠΙΛΟΓΗ ΔΕΙΓΜΑΤΟΣ ΚΑΙ ΠΟΙΟΤΙΚΗ
ΑΝΑΛΥΣΗ ΣΤΗΝ ΕΡΕΥΝΑ ΠΑΝΕΥΡΩΠΑΪΚΟΥ
ΕΝΔΙΑΦΕΡΟΝΤΟΣ EU-SILC.**

Εισηγήτρια: Αγγελική Καζατζή

Επιβλέπων: Δημήτρης Φουσκάκης, Επίκουρος Καθηγητής ΕΜΠ

Αθήνα, 2011

Στους γονείς μου

Εισαγωγή

Οι στατιστικές υπηρεσίες της Ευρωπαϊκής ένωσης ιδρύθηκαν με στόχο να επεξεργαστούν στατιστικά δεδομένα σε πανευρωπαϊκό επίπεδο και ακολούθως να εφοδιάσουν την Ευρωπαϊκή Ένωση με υψηλού επιπέδου στατιστικά αποτελέσματα, που αφορούν στο σύνολο των Ευρωπαϊκών χωρών. Κύριος στόχος των υπηρεσιών είναι η επίτευξη σύγκρισης στατιστικών δεδομένων μεταξύ χωρών ή περιοχών για ζητήματα πανευρωπαϊκού ενδιαφέροντος.

Η ύπαρξη αντικειμενικών και αξιόπιστων στατιστικών δεδομένων είναι αναγκαία για τους ανθρώπους, που παίρνουν τις αποφάσεις σε ζητήματα πανευρωπαϊκού ενδιαφέροντος, σε ζητήματα των κρατών μελών της Ευρωπαϊκής Ένωσης, των τοπικών αυτοδιοικήσεων αλλά και των επιχειρήσεων. Σημαντικότερη είναι επίσης η σημασία τους για το κοινό και τα Μέσα Μαζικής Ενημέρωσης ώστε να μπορεί να υπάρχει μία αξιόπιστη και αντικειμενική «εικόνα» πάνω στα φλέγοντα ζητήματα εποχής. Με την έννοια αυτή, κανείς μπορεί να ελέγξει και να συγκρίνει την πολιτική της χώρας του σε πανευρωπαϊκά ζητήματα με την πολιτική άλλων χωρών. Αυτό είναι πολύ σημαντικό γιατί δημιουργείται τρόπος αξιολόγησης μιας πολιτικής αλλά και των ανθρώπων που έχουν αποφασίσει για την πολιτική αυτή. Τέλος, τα αποτελέσματα από τη συλλογή δεδομένων σε πανευρωπαϊκό επίπεδο μπορούν να συμβάλλουν στην επίτευξη εθνικών σκοπών και στόχων.

Στα πλαίσια αυτά δρομολογούνται έρευνες, που αφορούν κοινωνικά, οικονομικά καθώς και πολιτικά ζητήματα έτσι ώστε μέσα από τα στατιστικά δεδομένα, που συγκεντρώνονται, να καλύπτονται οι τομείς και οι στόχοι, που προαναφέρθηκαν.

Μία από τις πανευρωπαϊκές αυτές έρευνες είναι η EU-SILC, η έρευνα που πρόκειται να ασχοληθούμε στη συνέχεια. Σκοπός της ακόλουθης εργασίας είναι η ανάλυση μίας ποιοτικής έκθεσης μέσα από την έρευνα EU-SILC. Μία ποιοτική έκθεση κατασκευάζεται και παραδίδεται από τις αρμόδιες στατιστικές υπηρεσίες ή από στατιστικές εταιρίες και γενικά αποτελεί το βασικό τμήμα ανάλυσης των ποιοτικών χαρακτηριστικών μίας έρευνας που έχει προηγηθεί.

Πιο αναλυτικά, στο Κεφάλαιο 1 κρίθηκε αναγκαία η εισαγωγή του αναγνώστη στη περιγραφή της έρευνας EU-SILC αλλά και των βασικών χαρακτηριστικών της. Στη συνέχεια, στο Κεφάλαιο 2, δεδομένου ότι η EU-SILC είναι μία δειγματοληπτική έρευνα, παρουσιάζονται η μεθοδολογία επιλογής του δείγματος, το κατάλληλο μέγεθος δείγματος αλλά γίνεται και μία ανάλυση της μεθοδολογίας των παραπάνω. Ακολούθως, στο Κεφάλαιο 3 παρουσιάζεται η μεθοδολογία κατασκευής μίας ποιοτικής έκθεσης. Πιο συγκεκριμένα δίνονται αναλυτικά τα τμήματα από τα οποία πρέπει να αποτελείται μία ποιοτική έκθεση αλλά και τα ποιοτικά χαρακτηριστικά της έρευνας που πρέπει να καταγράφονται σε κάθε τμήμα. Τέλος, στο Κεφάλαιο 4 παρουσιάζονται και αναλύονται τα δημοσιευμένα δεδομένα της EU-SILC, δεδομένα βέβαια που αποτελούν ποιοτικά χαρακτηριστικά και συμπεριλαμβάνονται σε μία ποιοτική έκθεση.

Το ζήτημα των ποιοτικών εκθέσεων είναι επίκαιρο και αυτός είναι και ο κύριος λόγος παρουσίας και ανάλυσης του.

Ευχαριστίες

Σε αυτό το σημείο, αισθάνομαι ειλικρινά υπόχρεη να αναφερθώ με μεγάλη χαρά στους ανθρώπους εκείνους που συντέλεσαν έμπρακτα, στην ολοκλήρωση αυτής της εργασίας. Ευχαριστώ θερμά τον καθηγητή μου κ. Φουσκάκη Δημήτριο, για την άψογη συνεργασία μας και την ενεργή βοήθειά του μέσα από τις διορθώσεις του και τις επισημάνσεις του καθ' όλη τη διάρκεια της συγγραφής αυτής. Ευχαριστώ θερμά επίσης, τους καθηγητές μου κ. Κοκολάκη Γεώργιο και κ. Σπηλιώτη Ιωάννη, που επέλεξαν να παρακολουθήσουν και να αξιολογήσουν την εργασία μου. Τέλος, θερμές ευχαριστίες οφείλω στην Κλειδέρη Μαρία, σύμβουλο στατιστικών μελετών, για τις πολύτιμες συμβουλές και οδηγίες της.

Περιεχόμενα:

Κεφάλαιο 1

Περιγραφή της έρευνας EU-SILC, ο σκοπός και οι μεταβλητές που μελετώνται

.....	- 7 -
1.1. Ευρωπαϊκή έρευνα για το εισόδημα και τις συνθήκες ζωής (European Union Statistics on Income and Living Conditions-EU-SILC)	- 7 -
1.2. Στόχοι και μεταβλητές που μελετώνται μέσα από την EU-SILC.....	- 8 -
1.2.1. Περιγραφή και διαφορά των δύο τύπων στατιστικών δεδομένων που συλλέγονται	- 9 -
1.2.2. Κοινό πλαίσιο εφαρμογής της EU-SILC	- 9 -
1.2.2.1. Βασικές και δευτερεύοντες Μεταβλητές που μελετώνται.....	- 10 -
1.2.2.2. Σχεδιασμός της έρευνας.....	- 11 -
1.2.2.3. Απαιτήσεις της EU-SILC.....	- 13 -
1.2.2.3. Απαιτήσεις της EU-SILC.....	- 13 -
1.2.2.4. Κοινές Έννοιες.....	- 14 -
1.2.2.5. Ταξινόμηση.....	- 15 -

Κεφάλαιο 2

Μεθοδολογία Επιλογής, Μέγεθος και Χαρακτηριστικά Δείγματος.....

.....	- 17 -
2.1. Επιλογή δείγματος νοικοκυριών	- 17 -
2.1.1. Κατασκευή δείγματος ατόμων από δείγμα νοικοκυριών.....	- 19 -
2.1.2. Δειγματοληψία περιοχής σε δύο στάδια που εφαρμόστηκε στην EU-SILC	- 20 -
2.2. Μέγεθος δείγματος	- 22 -
2.2.1. Απαιτήσεις μεγέθους του δείγματος	- 22 -
2.2.2. Η λογική επιλογής του ελαχίστου αποτελεσματικού μεγέθους δείγματος	- 26 -
2.2.2.1. Το τμήμα της EU-SILC που μελετά τις σύγχρονες μεταβλητές....	- 26 -
2.2.2.2. Το τμήμα της EU-SILC που μελετά τις διαχρονικές μεταβλητές..	- 29 -

Κεφάλαιο 3

Ποιοτική Έκθεση (Quality Report)

.....	- 32 -
3.1. Εισαγωγή στη στατιστική διαδικασία και τα αποτελέσματά της	- 34 -
3.2. Συνάφεια (Relevance).....	- 35 -
3.3. Ακρίβεια (Accuracy).....	- 37 -
3.3.1. Γενικό πλαίσιο	- 37 -
3.3.2. Αξιολόγηση.....	- 40 -
3.3.3. Δειγματοληπτικές έρευνες	- 41 -
3.3.3.1. Δειγματοληπτικά σφάλματα που βασίζονται σε τυχαία επιλογή δείγματος.....	- 42 -
3.3.3.2. Δειγματοληπτικά σφάλματα που δε βασίζονται σε τυχαίο δείγμα	- 43 -
3.4. Επικαιρότητα και εγκαίριότητα (Timeliness and punctuality)	- 44 -
3.5. Προσβασιμότητα και σαφήνεια (Accessibility and clarity).....	- 45 -
3.6. Συνοχή και συγκρισιμότητα (Coherence and comparability).....	- 46 -
3.6.1. Περιπτώσεις συνοχής/συγκρισιμότητας	- 47 -
3.6.2. Λόγοι έλλειψης συνοχής/συγκρισιμότητας	- 48 -
3.7. Συμβιβασμοί μεταξύ των ποιοτικών συνιστωσών των αποτελεσμάτων.	- 50 -
3.8. Εκτίμηση των αναγκών και των αντιλήψεων των χρηστών.	- 51 -
3.9. Επιδόσεις, κόστος και επιβάρυνση ανταπόκρισης.....	- 52 -

3.10. Εμπιστευτικότητα, Διαφάνεια και Ασφάλεια	- 54 -
3.11. Συμπεράσματα	- 56 -
Κεφάλαιο 4	
Πληροφορίες τμημάτων ποιοτικής έκθεσης της έρευνας EU-SILC	- 57 -
4.1. Εισαγωγή στη στατιστική διαδικασία και στα αποτελέσματά της.	- 57 -
4.1.1. Η στατιστική διαδικασία.....	- 57 -
4.1.2. Περιγραφή των συλλεγόμενων δεδομένων.....	- 58 -
4.2. Σύναφεια (Relevance).....	- 59 -
4.3. Ακρίβεια(Accuracy).....	- 59 -
4.4. Επικαιρότητα και εγκαιρότητα (Timeliness and punctuality)	- 60 -
4.5. Συγκρισιμότητα (Comparability).....	- 61 -
4.5.1 Ετήσιο Εισόδημα	- 61 -
4.5.2 Ποσοστό φτώχειας.....	- 66 -
Παράρτημα	- 72 -
Βιβλιογραφία	- 77 -

Κεφάλαιο 1

Περιγραφή της έρευνας EU-SILC, ο σκοπός και οι μεταβλητές που μελετώνται

1.1. Ευρωπαϊκή Έρευνα για το εισόδημα και τις συνθήκες ζωής (European Union Statistics on Income and Living Conditions-EU-SILC)

Λίγα λόγια για τη Eurostat

Η Eurostat ιδρύθηκε το 1953 και είναι μία από τις μεγαλύτερες στατιστικές υπηρεσίες την Ευρωπαϊκής Ένωσης. Ο ρόλος της Eurostat είναι να εφοδιάζει με στατιστικά αποτελέσματα την Ευρωπαϊκή Επιτροπή και άλλους ευρωπαϊκούς φορείς με δεδομένα έτσι ώστε να καθορίζουν, να εφαρμόζουν και να αναλύουν την Ευρωπαϊκή Πολιτική. Η Eurostat είναι η ευρωπαϊκή στατιστική υπηρεσία που δημοσιεύει και τα αποτελέσματα της EU-SILC.

Η ιδέα της Ευρωπαϊκής Έρευνας για το εισόδημα και τις συνθήκες ζωής των κατοίκων των χωρών της Ευρωπαϊκής Ένωσης επινοήθηκε το 2003 από τα κράτη-μέλη της Ευρωπαϊκής Ένωσης αλλά και την Ευρωπαϊκή Επιτροπή. Η έρευνα αυτή αποτελεί τη βασική πηγή για την ανάλυση της οικονομικής και κοινωνικής κατάστασης της Ευρωπαϊκής Ένωσης και των κρατών-μελών της, και επιπλέον συμβάλλει στην συγκριτική ανάλυση της κατανομής του εισοδήματος και των συνθηκών ζωής στην Ευρώπη. Η σύγκριση δύναται να γίνει τόσο μεταξύ των κρατών μελών της Ευρωπαϊκής Ένωσης αλλά και ορισμένων γειτονικών χωρών (Ισλανδία,, Νορβηγία, Ελβετία, Τουρκία).

Η εφαρμογή της έρευνας έγινε σταδιακά από τις χώρες από το 2004 και μετά. Μέχρι το 2010 έχουν λάβει μέρος στην έρευνα οι 27 χώρες κράτη μέλη της Ευρωπαϊκής Ένωσης καθώς επίσης η Ισλανδία, η Νορβηγία, η Ελβετία και η Τουρκία. Επίσης η έρευνα δοκιμάστηκε και στη συνέχεια εφαρμόστηκε από την Κροατία, τη Σερβία και την Πρώην Γιουγκοσλαβική Δημοκρατία της Μακεδονίας (ΠΓΔΜ). Αναλυτικά η εφαρμογή της έρευνας ξεκίνησε και συνεχίστηκε από τις χώρες ως εξής:

Πίνακας 1: Εφαρμογή της Έρευνας για το εισόδημα και τις συνθήκες ζωής (EU-SILC)

	2003	2004	2005	2006	2007	2008	2009	2010	2011
Βέλγιο									
Δανία									
Ιρλανδία									
Ελλάδα									
Λουξεμβούργο									
Αυστρία									
Νορβηγία									
Εσθονία									
Ισπανία									
Γαλλία									

Ιταλία									
Πορτογαλία									
Φιλανδία									
Σουηδία									
Ισλανδία									
Τσεχική Δημοκρατία									
Γερμανία									
Κύπρος									
Λετονία									
Λιθουανία									
Ουγγαρία									
Μάλτα									
Ολλανδία									
Πολωνία									
Σλοβενία									
Σλοβακία									
Ηνωμένο Βασίλειο									
Βουλγαρία									
Τουρκία			Tested						
Ρουμανία									
Ελβετία									
Κροατία						Tested	Tested	Tested	
Σερβία								Tested	
ΠΓΔΜ							Tested	Tested	Tested



Ολοκληρωμένη Εφαρμογή
Δοκιμαστική Εφαρμογή

1.2. Στόχοι και μεταβλητές που μελετώνται μέσα από την EU-SILC

Πρωταρχικός στόχος της εφαρμογής της έρευνας είναι να εδραιωθεί ένα **κοινό πλαίσιο** για τη συστηματική παραγωγή στατιστικών δεδομένων που αφορούν το εισόδημα και τις συνθήκες ζωής των κατοίκων της Ευρωπαϊκής Ένωσης. Στόχος είναι να παράγονται **σύγχρονα** και **διαχρονικά** δεδομένα για το εισόδημα, το επίπεδο και τη σύνθεση της φτώχειας καθώς και για τον κοινωνικό αποκλεισμό σε εθνικό και πανευρωπαϊκό επίπεδο. Όσον αφορά τη μεταβλητή του κοινωνικού αποκλεισμού η έρευνα EU-SILC μέσα από τα δεδομένα που συγκεντρώνει μελετά τις περιπτώσεις απόκλισης ενός ατόμου από το κοινωνικό σύνολο λόγω της οικονομικών, κοινωνικών ή άλλων παραγόντων. Η συνεργασία των κρατών-μελών με την Eurostat στοχεύει στη δυνατότητα σύγκρισης (comparability) των στατιστικών δεδομένων που συλλέγονται ανά χώρα καθώς και των αποτελεσμάτων που προκύπτουν μετά την επεξεργασία τους κάτι που θα επιδιωχθεί μέσα από την ανάπτυξη των μεθοδολογικών μελετών.

Ειδικότερα, η EU-SILC εδραιώθηκε με σκοπό να συλλέξει συγκρίσιμα, και πολυδιάστατα δεδομένα. Οι τομείς που στοχεύουν να καλύψουν τα συλλεχθέντα δεδομένα μέσα από τις **μεταβλητές** που μελετώνται αφορούν: α) το εισόδημα β) τη

φτώχεια γ) την κοινωνική απομόνωση δ) τη στέγαση ε) την εργασία στ) την εκπαίδευση και ζ) την υγεία.

1.2.1. Περιγραφή και Διαφορά των δύο τύπων στατιστικών δεδομένων που συλλέγονται

Δύο τύποι στατιστικών δεδομένων συλλέγονται ανά χρόνο από τις χώρες που εφαρμόζουν την έρευνα EU-SILC.

- **Σύγχρονα δεδομένα (Cross-sectional data):** Σύγχρονα δεδομένα ονομάζονται τα δεδομένα που αφορούν μια δεδομένη χρονική περίοδο ή ένα ορισμένο χρονικό διάστημα. Στην περίπτωση της Έρευνας μας, η χρονική αυτή περίοδος είναι είτε η τρέχουσα περίοδος ή κοντινή περίοδος σε αυτή που γίνεται η συλλογή των δεδομένων από το δείγμα. Πιο συγκεκριμένα, οι χώρες που εφαρμόζουν την έρευνα το χρόνο N συλλέγουν τα δεδομένα που αφορούν το χρόνο N και καλούνται να παραδώσουν μία ενδιάμεση **ποιοτική ανάλυση** στις αρμόδιες ευρωπαϊκές υπηρεσίες το Νοέμβριο του χρόνου N + 1. Η ποιοτική ανάλυση αφορά τα σύγχρονα δεδομένα που συλλέχθηκαν τη χρονιά N τα οποία περιέχουν μεταβλητές για την εξέταση του εισοδήματος, της φτώχειας, του κοινωνικού αποκλεισμού και άλλων συνθηκών ζωής. Στόχος των σύγχρονων δεδομένων είναι η εκτίμηση των επιπέδων των μεταβλητών που εξετάζονται αλλά και η εκτίμηση των αλλαγών από χρόνο σε χρόνο.
- **Διαχρονικά δεδομένα (Longitudinal data):** Τα διαχρονικά δεδομένα αφορούν σε ατομικό επίπεδο αλλαγές με την πάροδο του χρόνου και είναι στοιχεία που παρατηρούνται και συλλέγονται περιοδικά για τετραετή περίοδο. Τα διαχρονικά στοιχεία που συλλέγονται περιορίζονται σε πληροφορίες που αφορούν το εισόδημα, την εργασία καθώς και ένα περιορισμένο αριθμό από ποιοτικά κρίσιμες μεταβλητές στέρησης. Στόχος αυτών των δεδομένων της Έρευνας είναι ο εντοπισμός των επιπτώσεων και των δυναμικών διαδικασιών εμμονής στη φτώχεια, αλλά και ο κοινωνικός αποκλεισμός μεταξύ των υποομάδων του πληθυσμού. Οι χώρες καλούνται να παραδώσουν στις αρμόδιες ευρωπαϊκές υπηρεσίες μία ποιοτική ανάλυση των διαχρονικών δεδομένων που αφορούν την περίοδο μεταξύ της χρονιάς N-3 και της χρονιάς N, το Μάρτιο της χρονιάς N + 2.

Το μέγεθος του δείγματος που επιλέγεται για τα διαχρονικά δεδομένα είναι περιορισμένο σε σχέση με αυτό που επιλέγεται για τα σύγχρονα δεδομένα. Επίσης, τόσο τα σύγχρονα όσο και τα διαχρονικά δεδομένα ανανεώνονται κάθε χρόνο.

1.2.2. Κοινό πλαίσιο εφαρμογής της EU-SILC

Όλες οι χώρες της Ευρωπαϊκής Ένωσης καλούνται να εφαρμόσουν την έρευνα σε ένα κοινό πλαίσιο με στόχο την μεγιστοποίηση της δυνατότητας σύγκρισης

των παραχθέντων πληροφοριών μετά το τέλος την έρευνας. Οι χώρες που εφαρμόζουν την έρευνα καλούνται να έχουν εναρμονισμένες λίστες των βασικών (συλλέγονται κάθε χρόνο) αλλά και των δευτερευόντων **μεταβλητών** (συλλέγονται κάθε 4 ή περισσότερα χρόνια) που εξετάζονται, να εφαρμόζουν τον προτεινόμενο **σχεδιασμό** εφαρμογής της Έρευνας, να έχουν κοινές **απαιτήσεις** όσον αφορά τη συμπλήρωση των δεδομένων ή κάποιων πεδίων των δεδομένων που αγνοούνται, τον υπολογισμό των σφαλμάτων δειγματοληψίας, τη στάθμιση, κοινές **έννοιες** (νοικοκυριό, εισόδημα) και τέλος κοινή **ταξινόμηση** των συλλεχθέντων πληροφοριών (ISCO, NACE, ISCED). Αυτή είναι και η ουσιαστική διαφορά μεταξύ μίας κοινής έρευνας και μιας έρευνας με κοινό πλαίσιο. Άλλωστε το κοινό πλαίσιο προαπαιτείται από τις στατιστικές υπηρεσίες που συγκεντρώνουν τα δεδομένα από όλες τις χώρες και στη συνέχεια καλούνται να τα επεξεργαστούν και είναι αυτό που οδηγεί στην επιτυχημένη σύγκριση των συμπερασμάτων της έρευνας ανάμεσα στις χώρες. Όπως έχει προαναφερθεί ο κύριος στόχος των στατιστικών υπηρεσιών της Ευρώπης είναι η επίτευξη σύγκρισης στατιστικών δεδομένων μεταξύ χωρών ή περιοχών. Συνεπώς η εφαρμογή την Έρευνας σε ένα κοινό πλαίσιο είναι αναγκαία.

1.2.2.1. Βασικές και Δευτερεύοντες Μεταβλητές που μελετώνται

Η EU-SILC είναι πλέον μία πολυδιάστατη βάση δεδομένων. Τα δεδομένα εστιάζουν στο εισόδημα αλλά καλύπτουν και άλλους τομείς που αφορούν τη στέγαση, την εργασία, την υγεία, τη δημογραφία, την εκπαίδευση και τη στέρηση έτσι ώστε να μπορέσουν να μελετηθούν οι συνθήκες διαβίωσης και κατ' επέκταση ο κοινωνικός αποκλεισμός.

Οι βασικές μεταβλητές στοχεύουν σε πληροφορίες είτε για τα νοικοκυριά είτε σε πληροφορίες σε ατομικό επίπεδο (οι οποίες αφορούν αποκλειστικά άτομα με ηλικία άνω των 16 χρόνων). Ειδικότερα στο επίπεδο του νοικοκυριού συλλέγονται πληροφορίες για τους τομείς:

1. Βασικά Δεδομένα για το Νοικοκυριό
2. Εισόδημα
3. Στέγαση
4. Κοινωνικός Αποκλεισμός
5. Εργασία

Σε ατομικό επίπεδο συγκεντρώνονται πληροφορίες στους τομείς:

1. Βασικές Δημογραφικές πληροφορίες
2. Εισόδημα
3. Εκπαίδευση
4. Εργασία
5. Υγεία

Συνεπώς έχουμε μεταβλητές που μετρώνται σε επίπεδο νοικοκυριών αλλά και μεταβλητές που μετρώνται σε ατομικό επίπεδο. Μέσα από το δείγμα συλλέγονται πληροφορίες για το μέγεθος και τη σύνθεση των νοικοκυριών και των βασικών χαρακτηριστικών των μελών ενός νοικοκυριού, πληροφορίες για το εισόδημα και

άλλους τομείς σε ατομικό επίπεδο που βέβαια μετά μπορούν να υπολογιστούν και σε επίπεδο νοικοκυριού και τέλος πιο σύνθετες πληροφορίες που δεν αφορούν το εισόδημα αλλά αφορούν την κοινωνική κατάσταση και συλλέγονται επίσης σε ατομικό επίπεδο.

Οι δευτερεύοντες μεταβλητές που στοχεύει να μελετήσει η έρευνα εισάγονται στην έρευνα κάθε τέσσερα ή περισσότερα χρόνια και μόνο στο τμήμα της έρευνας που συλλέγει τα σύγχρονα δεδομένα. Από το 2005 ένα μοντέλο εισαγωγής μεταβλητών υπάρχει σύμφωνα με το οποίο έχουν εισαχθεί στην έρευνα ανά χρόνο οι παρακάτω μεταβλητές: α) Γενεαλογική μεταβίβαση της φτώχειας, η οποία αφορά τη μελέτη του κατά πόσο η φτώχεια μεταβιβάζεται από γενιά σε γενιά και εάν ναι σε τι βαθμό (2005), β) Κοινωνική Συμμετοχή (2006), γ) Συνθήκες Στέγασης (2007), δ) Υπερχρέωση και Οικονομικός Αποκλεισμός (2008), ε) Υλική Στέρηση (2009), στ) Ενδοθηλιακή κατανομή των πόρων, η οποία αφορά τη μελέτη του πώς μοιράζονται οι πόροι του νοικοκυριού στα μέλη του (2010). Τέλος 3 ακόμα μεταβλητές στοχεύεται να εισαχθούν στην έρευνα (ανά 1 χρόνο) και είναι ζ) η μελέτη της Διαγενεαλογικής μεταβίβασης των μειονεκτημάτων (2011), η) Συνθήκες Στέγασης (2012), θ) Ευημερία (2013)

1.2.2.2. Σχεδιασμός της έρευνας

Με στόχο τη θεμελίωση της Έρευνας για το Εισόδημα και τις συνθήκες ζωής στις Εθνικές Στατιστικές Υπηρεσίες των χωρών που συμμετέχουν στην έρευνα ο σχεδιασμός της έρευνας δεν είναι αυστηρός. Με την έννοια αυτή κάθε χώρα μπορεί να προσαρμόσει τα δεδομένα που χρειάζεται να συγκεντρωθούν για την EU-SILC σε έρευνες που ήδη υπάρχουν. Δηλαδή, παρόλο που τα διαχρονικά δεδομένα μπορούν, και συχνά γίνεται, να συνδεθούν με τα σύγχρονα δεδομένα αυτό δεν είναι απαραίτητο. Συνεπώς τα διαχρονικά και τα σύγχρονα δεδομένα μπορεί να προκύπτουν από διαφορετικές πηγές έτσι ώστε να εξυπηρετούνται οι εθνικές στατιστικές υπηρεσίες που άλλωστε είναι οι υπεύθυνες για τη συλλογή αυτών. Πιο συγκεκριμένα κάθε χώρα μπορεί να συλλέξει τα απαιτούμενα δεδομένα προς επεξεργασία μέσα από:

- Μία ή περισσότερες εθνικές πηγές. Οι πηγές αυτές μπορεί να είναι είτε άλλες έρευνες αλλά και τα μητρώα των χωρών.
- Μία ή περισσότερες εθνικές έρευνες συνδυασμένες με μία καινούρια έρευνα.
- Μία νέα έρευνα που να καλύπτει όλες τις απαιτήσεις της Έρευνας για το εισόδημα και τις συνθήκες ζωής.

Όσον αφορά τις χώρες που ξεκίνησαν νέα έρευνα για να καλύψουν τις απαιτήσεις της EU-SILC η στατιστική εταιρεία (Eurostat) που συγκεντρώνει τα δεδομένα όλων των χωρών πρότεινε έναν ολοκληρωμένο σχεδιασμό, τον λεγόμενο «εναλλασσόμενο σχεδιασμό» (rotational design) στις χώρες αυτές.

Εναλλασσόμενος σχεδιασμός

Με στόχο ο προτεινόμενος σχεδιασμός να ανταποκρίνεται τόσο στις απαιτήσεις της συλλογής των σύγχρονων δεδομένων όσο και στις απαιτήσεις της συλλογής των διαχρονικών δεδομένων προτάθηκε ο λεγόμενος «εναλλασσόμενος σχεδιασμός». Ο ολοκληρωμένος σχεδιασμός που προτάθηκε βασίζεται σε ένα αριθμό από υποδείγματα του πληθυσμού ή επαναλήμματα (replications)¹ τα οποία έχουν το ίδιο μέγεθος, τον ίδιο σχεδιασμό και είναι αντιπροσωπευτικά του πληθυσμού. Από τη μία χρονιά στην άλλη κάποια από αυτά μπορεί να διατηρούνται και άλλα να καταργούνται και να αντικαθιστώνται από άλλα. Ένας από τους πρωταρχικούς στόχους του τμήματος της EU-SILC που καλύπτει τα διαχρονικά δεδομένα είναι ο υπολογισμός του ποσοστού της μόνιμης απειλής της φτώχειας. Για να υπολογιστεί όμως το ποσοστό αυτό υπολογίζεται το ποσοστό των ανθρώπων που ζει σε νοικοκυριά με διαθέσιμο εισόδημα κάτω από το όριο του 60% για το τρέχον έτος της έρευνας αλλά και για τουλάχιστον 2 από τα προηγούμενα 3 έτη. Συνεπώς προκύπτει άμεσα ότι ο πληθυσμός που θα επιλεγεί την πρώτη χρονιά πρέπει να παρακολουθείται και για τα επόμενα 4 έτη τουλάχιστον. Αυτός είναι και ο κύριος λόγος που επιλέγονται 4 υποδείγματα ή επαναλήμματα. Βασικό χαρακτηριστικό του ολοκληρωμένου σχεδιασμού είναι ότι τόσο τα σύγχρονα όσο και τα διαχρονικά δεδομένα προκύπτουν από το ίδιο σει παρατηρήσεων του δείγματος και έτσι αποφεύγονται μη αναγκαίες επικαλύψεις όπου τελείως ξένες μεταξύ του έρευνες θα γινόντουσαν για τη συλλογή των σύγχρονων και διαχρονικών δεδομένων.

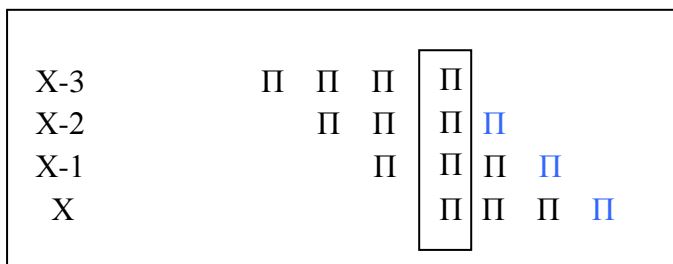
Μεθοδολογία εναλλασσόμενου σχεδιασμού

Θεωρούμε ότι τον πρώτο χρόνο εφαρμογής της έρευνας συγκεντρώνονται 4 πάνελ από δεδομένα (συμβ. Π). Την επόμενη χρονιά μία από αυτές τις ομάδες δεδομένων δε χρησιμοποιείται και στη θέση της μπαίνει μία νέα ομάδα δεδομένων που συλλέγεται τη τρέχουσα χρονιά (συμβ. Π). Η διαδικασία αυτή συνεχίζεται με τον ίδιο τρόπο κάθε χρόνο. Με αυτό το σχεδιασμό τόσο τα συμπεράσματα που αφορούν τη συλλογή διαχρονικών δεδομένων είναι εφικτά όσο και τα συμπεράσματα που αφορούν τη συλλογή σύγχρονων δεδομένων, δεδομένου ότι κάθε χρόνο ένα από τα πάνελ δεδομένων ανανεώνεται. Η διαδικασία σύμφωνα με την οποία συγκεντρώνονται τα δεδομένα γίνεται μέσα από την επιλογή δείγματος, η οποία θα αναλυθεί στο κεφάλαιο 2.

¹ Επανάλημμα είναι η επανάληψη μιας πειραματικής συνθήκης έτσι ώστε η μεταβλητότητα του φαινομένου που μελετάται να μπορεί να εκτιμηθεί. Φυσικά δε συγχέεται με την επανάληψη μίας μέτρησης.

Μία εικόνα του εναλλασσόμενου σχεδιασμού είναι η ακόλουθη (όπου με X συμβολίζεται το έτος που γίνεται η έρευνα και με X-i τα έτη που έχουν προηγηθεί). Τα σύγχρονα δεδομένα για κάθε έτος συγκεντρώνονται από τα 4 πάνελ που αντιστοιχούν για κάθε έτος ενώ τα διαχρονικά δεδομένα συγκεντρώνονται μέσα από τα δεδομένα για τα 4 προηγούμενα έτη, τα οποία όπως φαίνεται έχουν ανανεωθεί (μπλε χρώμα).

Εικόνα 1: Εναλλασσόμενος σχεδιασμός



1.2.2.3. Απαιτήσεις της EU-SILC

Καταλογισμός: Το κοινό πλαίσιο πάνω στο οποίο καλούνται να εφαρμόσουν οι χώρες την EU-SILC απαιτεί κοινή μεθοδολογία όσον αφορά τον καταλογισμό των πεδίων των δεδομένων που αγνοούνται. Καταλογισμός είναι η υποκατάσταση των τιμών που αγνοούνται στα δεδομένα που συλλέγει η κάθε χώρα. Η μεθοδολογία καταλογισμού που χρησιμοποιείται διαφέρει ανάλογα με το είδος της μη ανταπόκρισης του δείγματος.² Αφού εφαρμοστεί η διαδικασία του καταλογισμού τα δεδομένα είναι έτοιμα να παραδοθούν προς ανάλυση. Βέβαια, τα δεδομένα παραδίδονται μαζί με το βαθμό καταλογισμού που δέχθηκαν. Η κοινή μεθοδολογία που καλούνται να εφαρμόσουν οι χώρες όσον αφορά τον καταλογισμό βοηθά σε μία πιο ολοκληρωμένη και ομοιογενή εικόνα των πληροφοριών που παρέχονται από τη EUROSTAT.

Στάθμιση (weighting): Μία ακόμη απαίτηση του κοινού πλαισίου που ζητήθηκε να εφαρμοστεί η EU-SILC από τις χώρες ήταν η κοινή μεθοδολογία που έπρεπε να εφαρμόσουν οι χώρες κατά τη διαδικασία της στάθμισης. Στάθμιση είναι η διαδικασία κατά την οποία σε επιλεγμένα στοιχεία δίνεται περισσότερο “βάρος” ώστε να αυξηθεί ο βαθμός που αυτά επηρεάζουν το αποτέλεσμα. Η στάθμιση εφαρμόζεται στη Στατιστική με στόχο την αντιστάθμιση του βαθμού μεροληψίας. Η διαδικασία της στάθμισης εφαρμόζεται σε πολλά τμήματα της EU-SILC.

π.χ. Κατά τη διαδικασία επιλογής των ομάδων δεδομένων που περιγράφηκαν στον προτεινόμενο σχεδιασμό της έρευνας. Με στόχο να μην υπάρχει μεροληψία στην επιλογή των κατοίκων που θα ανήκουν στην ομάδα χρησιμοποιείται η μέθοδος της στάθμισης. Η διαδικασία της στάθμισης αποτελεί επίσης λύση για την αντιμετώπιση της μη ανταπόκρισης.

² Το ζήτημα αυτό θα συζητηθεί περαιτέρω στην ποιοτική ανάλυση που αφορά τη μη ανταπόκριση (non-response).

Οι κοινές μεθοδολογίες που απαιτήθηκαν σε ότι αφορά τόσο τη στάθμιση όσο και τον καταλογισμό κρίθηκαν αναγκαίες για την επεξεργασία δεδομένων και τη στατιστική ανάλυση αυτών.

1.2.2.4. Κοινές Έννοιες

Οι χώρες που εφαρμόζουν την EU-SILC καλούνται να έχουν κοινή έννοια του νοικοκυριού και του εισοδήματος.

Εισόδημα

Το συνολικό διαθέσιμο εισόδημα ενός νοικοκυριού υπολογίζεται ως το άθροισμα των εισοδημάτων όλων των μελών του νοικοκυριού προσαυξημένο με το εισόδημα που λαμβάνεται σε επίπεδο «νοικοκυριού». Το εισόδημα σε επίπεδο «νοικοκυριού» αφορά τυχόν εισοδήματα από επιδόματα, κέρδη τα οποία δεν αφορούν ένα άτομο του νοικοκυριού αλλά όλα τα άτομα του νοικοκυριού. Όσον αφορά τις πληροφορίες που αγνοούνται σε μεμονωμένα ερωτηματολόγια εφαρμόζεται η μέθοδος του καταλογισμού.

Το συνολικό διαθέσιμο εισόδημα περιλαμβάνει:

- Το εισόδημα από την εργασία, το οποίο περιλαμβάνει τους μισθούς των εργαζομένων καθώς και τα κέρδη από αυτοαπασχόληση.
- Ιδιωτικά εισοδήματα από επενδύσεις και ακίνητα.
- Μεταβιβάσεις μεταξύ νοικοκυριών.
- Όλες τις κοινωνικές μεταβιβάσεις που εισπράττονται σε μετρητά συμπεριλαμβανομένων των συντάξεων γήρατος.

Είναι σαφές λοιπόν ότι είναι αναγκαία η κοινή έννοια του εισοδήματος αφού αν κάθε χώρα συμπεριλάμβανε στο συνολικό εισόδημα διαφορετικές συνιστώσες τότε θα υπήρχε κακή διαχείριση της πληροφορίας και σίγουρα της συμπερασματολογίας για τη μεταβλητή του εισοδήματος. Δεδομένου επίσης ότι οι πληροφορίες που συλλέγονται για τη μεταβλητή του εισοδήματος χρησιμοποιούνται περαιτέρω για την ανάλυση των συνθηκών ζωής είναι προφανές ότι είναι αναγκαία η αυστηρή απαίτηση για κοινή έννοια του εισοδήματος.

Σε αυτό το σημείο είναι αναγκαίο να αναφέρουμε την «κλίμακα ισοδυναμίας» που πρότεινε η Eurostat στις χώρες έτσι ώστε να ληφθεί υπόψη ο αντίκτυπος των διαφορών των νοικοκυριών σε μέγεθος και σύνθεση. Μπορούμε να πούμε η «κλίμακα ισοδυναμίας» είναι μία μορφή εφαρμογής της στάθμισης. Η κλίμακα ισοδυναμίας που πρότεινε η Eurostat είναι να συμπεριληφθεί ένας παράγοντας της τάξεως:

- 1,0 για το πρώτο μέλος της οικογενείας που είναι άνω των 14 ετών.
- 0,5 για τα υπόλοιπα μέλη της οικογενείας που είναι άνω των 14 ετών.
- 0,3 για τα μέλη της οικογενείας που έχουν ηλικία από 0-13 ετών.

Με τον ορισμό της κλίμακας ισοδυναμίας υπάρχει διαφοροποίηση στα νοικοκυριά που έχουν το ίδιο εισόδημα αλλά διαφέρουν στο μέγεθος και τη σύνθεσή τους. Είναι γεγονός ότι οι ανάγκες ενός νοικοκυριού διαφέρουν ανάλογα με το μέγεθος του νοικοκυριού και την ηλικία των μελών που το απαρτίζουν. Συνεπώς δε

θα μπορούσαμε να υπολογίσουμε το εισόδημα ενός νοικοκυριού ανεξάρτητα από τη σύνθεση και το μέγεθός του και στη συνέχεια να βγάλουμε συμπεράσματα από αυτό.

Νοικοκυριά

Η Eurostat πρότεινε μία ταξινόμηση των νοικοκυριών ανάλογα με τον αριθμό των ενηλίκων μελών, την ηλικία τους, και τον αριθμό των εξαρτώμενων παιδιών που ζουν μαζί τους. Με τη κοινή ταξινόμηση αυτή, κατά την επεξεργασία των δεδομένων, και όχι μόνο, είναι άμεση η αντίληψη της σύνθεσης και του μεγέθους ενός νοικοκυριού.

Κάθε νοικοκυριό ανήκει σε μία από τις παρακάτω κατηγορίες:

- ⇒ Νοικοκυριά χωρίς εξαρτώμενα παιδιά.
- ⇒ Νοικοκυριά με ένα άτομο.
 - ⇒ Έναν ενήλικα αρσενικού γένους.
 - ⇒ Έναν ενήλικα θηλυκού γένους.
 - ⇒ Έναν ενήλικα ηλικίας 65 ετών και άνω.
 - ⇒ Έναν ενήλικα ηλικίας μικρότερης των 65 ετών.
- ⇒ Δύο ενήλικες, χωρίς εξαρτώμενα παιδιά, νεότερους από 65 ετών.
- ⇒ Δύο ενήλικες, χωρίς εξαρτώμενα παιδιά, με τον ένα από αυτούς σε ηλικία 65 ετών και άνω.
- ⇒ Τρεις ή περισσότεροι ενήλικες, χωρίς εξαρτώμενα παιδιά.
- ⇒ Νοικοκυριά με εξαρτώμενα παιδιά.
 - ⇒ Ένας γονέας με τουλάχιστον ένα εξαρτώμενο παιδί.
 - ⇒ Δύο γονείς με 1 εξαρτώμενο παιδί.
 - ⇒ Δύο γονείς με 2 εξαρτώμενα παιδιά.
 - ⇒ Δύο γονείς με 3 ή περισσότερα εξαρτώμενα παιδιά.
 - ⇒ Τρεις ή περισσότεροι ενήλικες με εξαρτώμενα παιδιά.

Εξαρτώμενα παιδιά θεωρούνται όλα τα άτομα του νοικοκυριού με ηλικία κάτω των 18 ετών καθώς και τα οικονομικά ανενεργά άτομα ηλικίας 18 έως 24 ετών που ζουν με έναν τουλάχιστον από τους γονείς τους.

1.2.2.5. Ταξινόμηση

Το κοινό πλαίσιο σύμφωνα με το οποίο καλούνται να εφαρμόσουν οι χώρες την EU-SILC συμπεριλαμβάνει εκτός των άλλων κοινή ταξινόμηση της πληροφορίας που συλλέγει η κάθε χώρα. Η ταξινόμηση γίνεται σύμφωνα με το Διεθνές Πρότυπο Ταξινόμησης των Επαγγελματιών (ISCO-International Standard Classification of Occupations)³, τη Στατιστική Ταξινόμηση των Οικονομικών Δραστηριοτήτων (NACE- *Nomenclature statistique des Activités économiques dans la Communauté*

³http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/documents/tab/Tab/ISCO.pdf

Européenne)⁴ και το Διεθνές Πρότυπο Ταξινόμησης της Εκπαίδευσης (ISCED-International Standard Classification Of Education)⁵.

ISCO: Το Διεθνές Πρότυπο Ταξινόμησης των Επαγγελμάτων είναι ένα εργαλείο οργάνωσης των επαγγελμάτων σε ένα αυστηρά καθορισμένο σύνολο ομάδων. Τα επαγγέλματα ταξινομούνται στις ομάδες σύμφωνα με τα καθήκοντα και τις υποχρεώσεις που αναλαμβάνονται κατά την άσκηση του επαγγέλματος.

NACE: Η Στατιστική Ταξινόμηση των Οικονομικών Δραστηριοτήτων ορίζει την ονοματολογία των οικονομικών δραστηριοτήτων στην Ευρωπαϊκή Ένωση. Πρόκειται για μία τετραψήφια ταξινόμηση η οποία προσφέρει το πλαίσιο για τη συλλογή και την παρουσίαση ενός μεγάλου εύρους στατιστικών δεδομένων που αφορούν οικονομικές δραστηριότητες.

ISCED: Το Διεθνές Πρότυπο Ταξινόμησης της Εκπαίδευσης σχεδιάστηκε από την UNESCO και χρησιμεύει ως κατάλληλο μέσο για τη συλλογή, συγκέντρωση και παρουσίαση στατιστικών δεδομένων που αφορούν το τομέα της εκπαίδευσης, τόσο σε επίπεδο χωρών αλλά και διεθνώς.

Και τα τρία Πρότυπα ανανεώνονται και αναθεωρούνται τακτικώς. Σύμφωνα με τα τρία αυτά Πρότυπα έγινε η ταξινόμηση και κωδικοποίηση των δεδομένων που συλλέχθηκαν από τις χώρες που εφάρμοσαν την EU-SILC.

⁴http://epp.eurostat.ec.europa.eu/portal/page/portal/income_social_inclusion_living_conditions/documents/tab/Tab/NACE.pdf

⁵ http://www.uis.unesco.org/TEMPLATE/pdf/isced/ISCED_A.pdf

Κεφάλαιο 2

Μεθοδολογία Επιλογής, Μέγεθος και Χαρακτηριστικά Δείγματος

Όπως έχει αναφερθεί οι χώρες που λαμβάνουν μέρος στην EU-SILC είναι ελεύθερες ως ένα βαθμό να επιλέξουν τον σχεδιασμό της έρευνας. Παρόλα αυτά, το δειγματοληπτικό πλαίσιο και η επιλογή του δείγματος πρέπει να διασφαλίζει ότι κάθε πολίτης και κάθε νοικοκυριό του πληθυσμού έχει μία γνωστή και μη μηδενική πιθανότητα επιλογής στο δείγμα. Αυτό άλλωστε θέλουμε να ισχύει έτσι ώστε να μην εμφανίζεται **μεροληψία** ως προς την επιλογή του δείγματος και κατ' επέκταση το δείγμα να είναι **αντιπροσωπευτικό** του πληθυσμού.

Στις ποιοτικές αναλύσεις των χωρών που εφαρμόζουν την EU-SILC αναφέρεται ότι τα δείγματα που χρησιμοποιούνται είναι δείγματα πιθανότητας από το σχεδιασμό τους. Μία μέθοδος δείγματος πιθανότητας είναι οποιαδήποτε μέθοδος χρησιμοποιείται έτσι ώστε να διασφαλιστεί ότι τα διαφορετικά μέλη του πληθυσμού έχουν την ίδια πιθανότητα να εμφανιστούν στο δείγμα. Το γεγονός αυτό διασφαλίζει την προϋπόθεση του αντιπροσωπευτικού δείγματος από τον πληθυσμό. Στις περισσότερες χώρες πλέον υπάρχει ενημερωμένη λίστα η οποία είναι διαθέσιμη για την επιλογή του δείγματος. Πρακτικά, σε όλες τις χώρες χρησιμοποιήθηκε ένα άμεσα επιλεγόμενο δείγμα από νοικοκυριά/άτομα ή μία επιλογή δείγματος βασισμένη στις περιοχές σε δύο στάδια (two-stage area sampling).

Άμεση δειγματοληψία (direct sampling): Ο όρος άμεση δειγματοληψία χρησιμοποιείται όταν οι μονάδες του δείγματος είναι τα πραγματικά μέλη του πληθυσμού που μελετάται. Σε αντίθεση με την άμεση δειγματοληψία αρκετές φορές συναντάμε ως μονάδα του δείγματος έναν αριθμό. π.χ. αριθμό απογραφής. Ο όρος άμεση δειγματοληψία συνδέεται με την αμεσότητα της παρατήρησης των μονάδων του δείγματος και όχι με τη διαδικασία επιλογής τους.

Δειγματοληψία σε δύο στάδια (two-stage sample): Η δειγματοληψία σε δύο στάδια είναι η πιο απλή μορφή μιας δειγματοληψίας σε πολλαπλά στάδια. Στην περίπτωση της δειγματοληψίας σε δύο στάδια, ο πληθυσμός που θέλουμε να μελετηθεί χωρίζεται σε βασικές κατηγορίες καθεμία από τις οποίες περιέχει τις δευτερεύοντες μονάδες, που είναι και αυτές που θέλουμε να ανήκουν στο δείγμα. Στη συνέχεια, επιλέγεται ένα δείγμα από τις βασικές κατηγορίες από τις οποίες επιλέγεται ένα υπο-δείγμα το οποίο περιέχει και τις δευτερεύοντες μονάδες.

2.1. Επιλογή δείγματος νοικοκυριών

Για την πλειονότητα των χωρών που εφαρμόζουν την EU-SILC, οι στόχοι της Έρευνας εξυπηρετούνται καλύτερα από ένα δείγμα πιθανότητας νοικοκυριών. Όλα τα μέλη των νοικοκυριών που θα επιλεγθούν στο δείγμα θα ερωτηθούν έτσι ώστε να συλλεχθούν δεδομένα για τις μεταβλητές σε ατομικό επίπεδο. Συνεπώς από το δείγμα

πιθανότητας των νοικοκυριών θα προκύψει δείγμα ίσης πιθανότητας επιλογής και για τους κατοίκους της χώρας. Η επιλογή δείγματος νοικοκυριών είναι ιδανική για τη μελέτη των σύγχρονων μεταβλητών. Επίσης, είναι και η μέθοδος για την επιλογή του αρχικού δείγματος στην περίπτωση των διαχρονικών μεταβλητών. Όπως έχει προαναφερθεί για τη μελέτη των διαχρονικών μεταβλητών χρησιμοποιείται ο εναλλασσόμενος σχεδιασμός, συνεπώς ένα αρχικό δείγμα είναι απαραίτητο, το οποίο ανανεώνεται κάθε χρόνο. Βέβαια, ο μηχανισμός επιλογής δείγματος δύναται να διαφέρει αφού μπορεί να βασίζεται εναλλακτικά στην άμεση επιλογή δείγματος νοικοκυριών ή στην επιλογή δείγματος νοικοκυριών βάσει διευθύνσεων ή ατόμων.

Επιλογή δείγματος νοικοκυριών βάσει διευθύνσεων: Στην περίπτωση αυτή εφαρμόζεται ένα δείγμα πιθανότητας το οποίο όμως περιέχει διευθύνσεις και από το οποίο καταλήγουμε σε ένα δείγμα νοικοκυριών. Συνήθως υπάρχει μία “1-1” αντιστοιχία μεταξύ των δύο αυτών παραγόντων (διευθύνσεις – νοικοκυριά). Ομοίως βέβαια προκύπτει το δείγμα νοικοκυριών αν από το δείγμα διευθύνσεων που έχει επιλεγεί συμπεριλάβουμε στο δείγμα νοικοκυριών όλα τα νοικοκυριά που ανήκουν στη διεύθυνση που έχει επιλεγεί στο δείγμα διευθύνσεων. Σε περιπτώσεις συστηματικής εμφάνισης τέτοιων διευθύνσεων στο δείγμα διευθύνσεων, συχνά εμφανίζονται τεχνικές οι οποίες περιορίζουν το μέγιστο αριθμό νοικοκυριών που θα συμπεριληφθούν στο δείγμα νοικοκυριών, και έχουν την ίδια διεύθυνση. Οι τεχνικές αυτές οφείλουν να προλαμβάνουν οποιαδήποτε αλλαγή στην πιθανότητα επιλογής ενός νοικοκυριού (μείωση ή αύξηση).

Επιλογή δείγματος νοικοκυριών βάσει κατοίκων: Η επιλογή δείγματος νοικοκυριών γίνεται από μία λίστα ατόμων που ανήκουν σε δείγμα ίσης πιθανότητας ατόμων που συλλέγεται. Ουσιαστικά, ένα νοικοκυριό θα συμπεριληφθεί στο δείγμα νοικοκυριών εφόσον ένα τουλάχιστον μέλος του εμφανίζεται στο δείγμα ατόμων που συλλέχθηκε. Τα άτομα που μπορούν να συμπεριληφθούν στο δείγμα ατόμων είναι τα άτομα ηλικίας 16 ετών τουλάχιστον. Αυτό οδηγεί σε μία μη μηδενική πιθανότητα εμφάνισης ενός νοικοκυριού στο δείγμα εφόσον έχει τουλάχιστον ένα τέτοιο μέλος. Συνεπώς, από ένα δείγμα πιθανότητας ατόμων κατασκευάζεται ένα δείγμα πιθανότητας νοικοκυριών. Παρ’ όλα αυτά, από τη διαδικασία αυτή προκύπτουν 2 ζητήματα:

α) Στην περίπτωση που σε ένα νοικοκυριό κανένα μέλος δεν είναι άνω των 16 ετών η πιθανότητα να συμπεριληφθεί το νοικοκυριό αυτό στο δείγμα είναι μηδενική.

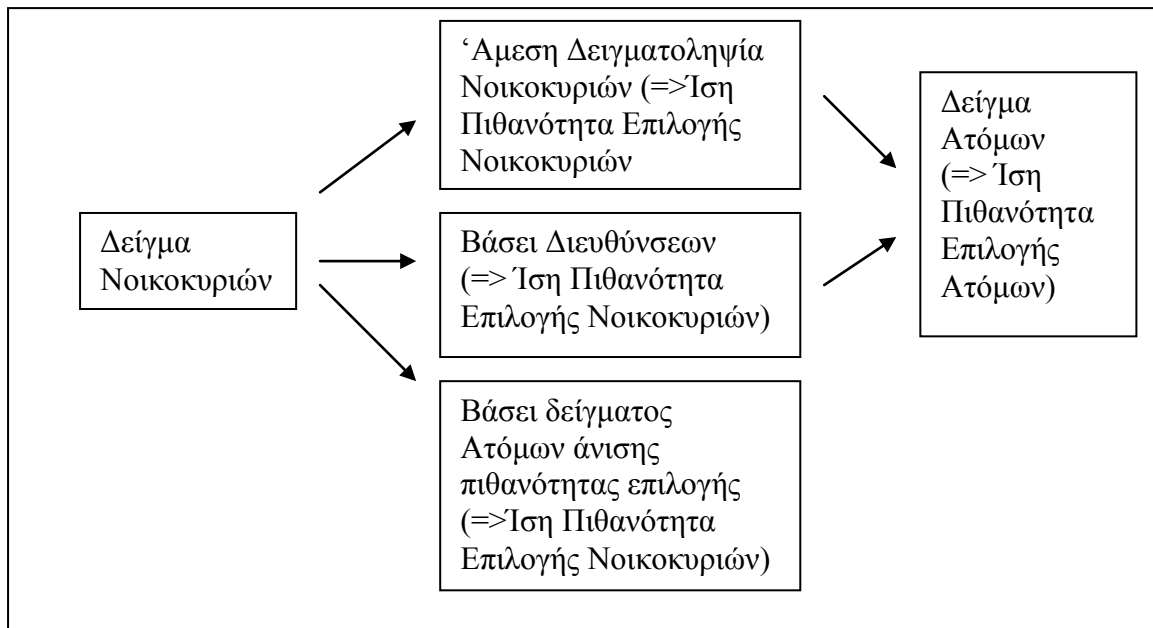
β) Εφόσον τα άτομα έχουν επιλεγεί για το δείγμα ατόμων με ίσες πιθανότητες επιλογής τότε όσο μεγαλύτερος είναι ο αριθμός των μελών, ηλικίας άνω των 16 ετών, ενός νοικοκυριού τόσο μεγαλύτερη είναι και η πιθανότητα να συμπεριληφθεί το νοικοκυριό στο δείγμα. Η αντιμετώπιση αυτού του ζητήματος μπορεί να γίνει με την μείωση της πιθανότητας επιλογής ενός ατόμου στο δείγμα εφόσον υπάρχει τουλάχιστον ακόμη ένα άτομο στο νοικοκυριό του, το οποίο ανήκει και στο δείγμα ατόμων. Η μείωση είναι ανάλογη του αριθμού των επιπλέον ατόμων που ανήκουν στο ίδιο νοικοκυριό αλλά και στο δείγμα ατόμων. Για παράδειγμα αν ένα άτομο ανήκει σε ένα νοικοκυριό στο οποίο ακόμη τρία άτομα του νοικοκυριού είναι άνω των 16 ετών

επιβάλλεται στο άτομο αυτό μία πιθανότητα επιλογής κατά 75% μικρότερη από την πιθανότητα επιλογής ενός ατόμου που αντιπροσωπεύει ένα νοικοκυριό για το οποίο δεν υπάρχει άλλο μέλος ηλικίας άνω των 16 ετών. Κατά συνέπεια τα δύο νοικοκυριά θα έχουν την ίδια πιθανότητα επιλογής στο δείγμα νοικοκυριών, τελικά. Εναλλακτικά, από το δείγμα των ατόμων που θα δημιουργηθεί συλλέγονται οι πληροφορίες για το νοικοκυριό στο οποίο ανήκουν και στη συνέχεια κατασκευάζεται ένα υπο-δείγμα στο οποίο η πιθανότητα επανεπιλογής ενός ατόμου είναι αντιστρόφως ανάλογη με τα μέλη των ατόμων που αντιπροσωπεύουν το νοικοκυριό στο οποίο ανήκει. Φυσικά το αρχικό δείγμα ατόμων που θα επιλεγεί πρέπει να είναι μεγαλύτερο από το μέγεθος του τελικού δείγματος νοικοκυριών που επιζητούμε. Παρ' όλα αυτά, οι δύο παραπάνω μέθοδοι απαιτούν την εύρεση πληροφοριών για το νοικοκυριό που αντιπροσωπεύει κάθε άτομο που επιλέγεται στο δείγμα. Στην περίπτωση που υπάρχουν οι πληροφορίες αυτές στα μητρώα των χωρών ή κάπου αλλού καταγεγραμμένα η διαδικασία αυτή δεν αυξάνει το κόστος (χρονικό, υλικό, οικονομικό). Σε αντίθετη περίπτωση η διαδικασία εύρεσης των πληροφοριών αυτών θα αυξήσει ουσιαστικά το κόστος αφού θα πρέπει να συλλεχθούν για το μεγάλο δείγμα των ατόμων που θα επιλεγεί.

2.1.1. Κατασκευή δείγματος ατόμων από δείγμα νοικοκυριών

Στο σημείο αυτό πρέπει να τονίσουμε ότι από οποιοδήποτε τρόπο επιλογής δείγματος νοικοκυριού μπορούμε να καταλήξουμε άμεσα ή έμμεσα σε ένα δείγμα ατόμων. Συγκεκριμένα, σε περίπτωση που θέλουμε να συλλέξουμε πληροφορίες από όλα τα μέλη ενός νοικοκυριού κατασκευάζεται ένα δείγμα στο οποίο ανήκουν όλα τα μέλη από κάθε νοικοκυριό που επιλέχθηκε στο αρχικό δείγμα νοικοκυριών. Από την άλλη πλευρά στην περίπτωση που θέλουμε ένα δείγμα ατόμων βάσει του δείγματος νοικοκυριών τότε από κάθε νοικοκυριό που ανήκει στο αρχικό δείγμα επιλέγεται τυχαία ένα μέλος που με τη σειρά του ανήκει στο δείγμα ατόμων ή επιλέγεται το μέλος που οδήγησε στην επιλογή του νοικοκυριού με τη μέθοδο επιλογής νοικοκυριών βάσει ατόμων. Βέβαια, αν θέλουμε ένα δείγμα ατόμων με ίση πιθανότητα επιλογής τότε εφ' όσον η δειγματοληψία των νοικοκυριών είναι άμεση ή βασίζεται στις διευθύνσεις η πιθανότητα επιλογής ενός ατόμου στο δείγμα ατόμων που θέλουμε να κατασκευάσουμε εξαρτάται από το μέγεθος του νοικοκυριού.

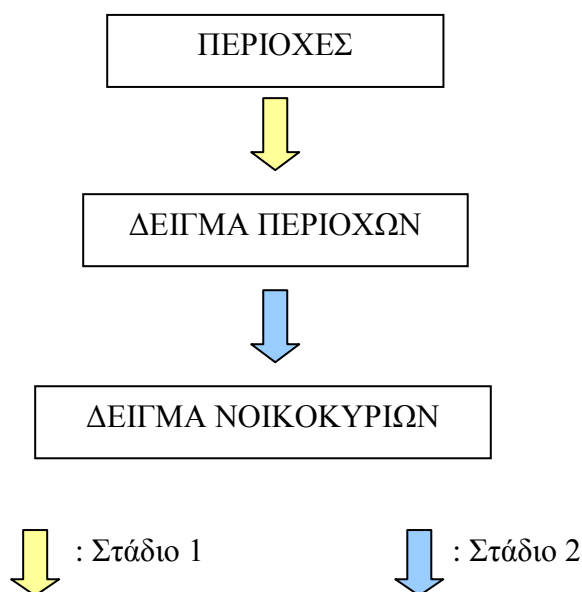
Εικόνα 2: Διάγραμμα επιλογής δείγματος



2.1.2. Δειγματοληψία περιοχής σε δύο στάδια που εφαρμόστηκε στην EU-SILC

Το δείγμα των νοικοκυριών επιλέχθηκε σε δύο στάδια. Όπως αναφέρθηκε στην περιγραφή μίας δειγματοληψίας σε δύο στάδια, έχουμε τις βασικές μονάδες δείγματος και τις δευτερεύοντες μονάδες δείγματος. Στην EU-SILC οι βασικές μονάδες δείγματος, δηλαδή οι βασικές κατηγορίες στις οποίες χωρίζεται ο πληθυσμός, είναι οι περιοχές. Περιοχές θεωρήθηκαν ένα ή περισσότερα οικοδομικά τετράγωνα. Δευτερεύοντες μονάδες δείγματος οι οποίες υπο-επιλέχθηκαν από το δείγμα περιοχών είναι τα νοικοκυριά. Στο **πρώτο στάδιο** δειγματοληψίας ανήκει η επιλογή ενός δείγματος περιοχών από το σύνολο των περιοχών στο οποίο χωρίστηκε ο πληθυσμός. Στο **δεύτερο στάδιο** δειγματοληψίας ανήκει η επιλογή ενός δείγματος από νοικοκυριά από κάθε περιοχή η οποία επιλέχθηκε κατά το πρώτο στάδιο δειγματοληψίας.

Εικόνα 3: Δειγματοληψία σε δύο στάδια



Κριτήρια στρωματοποίησης και υπο-στρωματοποίησης

Η διαδικασία χωρισμού του πληθυσμού σε περιοχές γίνεται σε δύο επίπεδα. Κατά το πρώτο επίπεδο έχουμε την γεωγραφική στρωματοποίηση η οποία βασίζεται στο διαχωρισμό της έκτασης της κάθε χώρας σε περιοχές. Η γεωγραφική διαστρωμάτωση γίνεται σύμφωνα με την Ευρωπαϊκή Ταξινόμηση NUTS II (Nomenclature of Territorial Units for Statistics, by regional level) έτσι ώστε να είναι κοινή για όλες τις χώρες που εφαρμόζουν την EU-SILC⁶. Σύμφωνα με τη NUTS II η έκταση κάθε χώρας χωρίζεται σε περιοχές και κάθε περιοχή κωδικοποιείται. Το δεύτερο επίπεδο στρωματοποίησης του πληθυσμού περιέχει την ομαδοποίηση δήμων και κοινοτήτων μέσα σε κάθε περιοχή που ορίστηκε προηγουμένως, σύμφωνα με το βαθμό αστικοποίησης (π.χ. τον πληθυσμό).

Εφαρμογή της δειγματοληψίας περιοχής σε δύο στάδια στην Ελλάδα:

Η χώρα κατά το πρώτο επίπεδο στρωματοποίησης χωρίστηκε σε 13 περιοχές σύμφωνα με την Ευρωπαϊκή Ταξινόμηση NUTS II. Αναφέρεται ότι όσον αφορά τα δύο πολεοδομικά μεγάλα συγκροτήματα της Αθήνας, και της Θεσσαλονίκης αποτελούν ξεχωριστά μεγάλα γεωγραφικά στρώματα.

Κατά το δεύτερο επίπεδο στρωματοποίησης η ομαδοποίηση των περιοχών σε δήμους και κοινότητες έγινε ανάλογα με τον αριθμό των κατοίκων. Πιο συγκεκριμένα υπήρξαν 4 ομάδες:

- ≥ 30.000 κατοίκους.
- 5.000 – 29.999 κατοίκους.
- 1.000 – 4.999 κατοίκους.
- 0- 999 κατοίκους.

⁶ file:///C:/DOCUME~1/angel/LOCALS~1/Temp/NUTS_33_07-04-2011_08-56-30.htm

Ο αριθμός των τελικών στρωμάτων στις 13 διοικητικές περιφέρειες ήταν 50. Όσον αφορά το γεωγραφικό στρώμα της Αθήνας, χωρίστηκε σε 31 τμήματα του ίδιου μεγέθους (το μέγεθος τους υπολογίζεται από τον αριθμό των νοικοκυριών που ανήκουν σε αυτό). Ο διαχωρισμός έγινε βάσει των δήμων που απαρτίζουν την περιοχή της Αθήνας αλλά και βάσει κοινωνικο-οικονομικών κριτηρίων. Κατά τον ίδιο τρόπο χωρίστηκε η περιοχή της Θεσσαλονίκης σε 9 ίδιου μεγέθους τμήματα.

Οι γεωγραφικές περιφέρειες των Αθηνών και της Θεσσαλονίκης καλύπτουν το 40% του πληθυσμού αλλά και ακόμη μεγαλύτερα ποσοστά σε κοινωνικο-οικονομικές μεταβλητές που εξετάζονται μέσα από την EU-SILC. Αυτό είχε σαν αποτέλεσμα να δημιουργηθούν 90 τελικά στρώματα για την εφαρμογή της έρευνας στην Ελλάδα.

Είναι φανερό λοιπόν ότι η μελέτη και η σωστή διαστρωμάτωση προλαμβάνει τη συλλογή μη αντιπροσωπευτικών δεδομένων. Ας φέρουμε ένα παράδειγμα για την περιοχή της Ελλάδας:

- ✓ Αν το δείγμα του πληθυσμού βασιζόταν σε τυχαίο δείγμα από το σύνολο του πληθυσμού της Ελλάδος, οι κάτοικοι της Αθήνας θα είχαν σαφώς μεγαλύτερη πιθανότητα να εμφανιστούν στο δείγμα αφού είναι σαφώς περισσότεροι από κατοίκους άλλων περιοχών. Το αποτέλεσμα θα ήταν ένα δείγμα του πληθυσμού από περισσότερους κατοίκους της Αθήνας και λιγότερους κατοίκους άλλων περιοχών. Αυτό βέβαια θα ήταν μη αντιπροσωπευτικό δείγμα του συνόλου του πληθυσμού της Ελλάδος, αφού τόσο οι συνθήκες ζωής όσο και το εισόδημα διαφέρουν.
- ✓ Επίσης στην περιοχή των Αθηνών υπάρχει μεγάλη διαφορά στις συνθήκες ζωής και το εισόδημα ανάμεσα στους δήμους. Ο περαιτέρω διαχωρισμός σε 31 τμήματα βοήθησε στη συλλογή αντιπροσωπευτικού δείγματος από την περιοχή της Αθήνας.

Συνεπώς, η διαδικασία χωρισμού της έκτασης των χώρας σε στρώματα κατά την επιλογή δείγματος κρίθηκε αναγκαία. Το τελικό δείγμα προκύπτει μετά την ολοκλήρωση της στρωματοποίησης και υπο-στρωματοποίησης της έκτασης της χώρας. Από κάθε περιοχή επιλέγεται τυχαίο δείγμα και το σύνολο των υπο-δειγμάτων αποτελεί το δείγμα του πληθυσμού της χώρας που θα ερωτηθεί για την εφαρμογή της EU-SILC.

2.2. Μέγεθος δείγματος

2.2.1. Απαιτήσεις μεγέθους του δείγματος

Το μέγεθος του δείγματος μιας έρευνας εξαρτάται από τα χαρακτηριστικά της έρευνας για την οποία κατασκευάζεται. Με την έννοια χαρακτηριστικά εννοούνται:

α) οι ουσιαστικές απαιτήσεις μίας έρευνας που εφαρμόζεται (σκοπός των πληροφοριών που πρόκειται να συλλεχθούν, απαιτήσεις ακρίβειας, ανάλυση των αποτελεσμάτων), β) ο παράγοντας του κόστους μιας έρευνας (προϋπολογισμός, τεχνικούς πόρους, response burden (=η προσπάθεια που απαιτείται συνολικά για την

ανταπόκριση στην έρευνα) και γ) οι πρακτικοί λόγοι (σκοπιμότητα, βιωσιμότητα, έλεγχο ποιότητας).

Όσον αφορά το μέγεθος του δείγματος της EU-SILC, αυτό εξαρτάται τόσο από τα προαναφερθέντα στοιχεία αλλά και από δ) επιπρόσθετα χαρακτηριστικά που αφορούν την εφαρμογή της EU-SILC. Όπως έχει γίνει ξεκάθαρο, οι πληροφορίες που συλλέγονται από την εφαρμογή της EU-SILC χρησιμοποιούνται για ανάλυση σε εθνικό επίπεδο αλλά και συγκριτική ανάλυση μεταξύ των χωρών που την εφαρμόζουν.

Το σύνολο των παραπάνω χαρακτηριστικών σε συνδυασμό με το ότι πρέπει να αποφασιστεί ένα ενιαίο μέγεθος δείγματος για όλες τις χώρες που εφαρμόζουν την έρευνα κάνει την απόφαση για το μέγεθος του δείγματος ένα πολυδιάστατο ζήτημα. Η **λογική** που κρύβεται πίσω από την επιλογή του μεγέθους των δειγμάτων θα συζητηθεί αναλυτικά τόσο για το διαχρονικό τμήμα όσο και για το σύγχρονο τμήμα της EU-SILC. Όπως έχει συζητηθεί οι μεταβλητές που μελετά η EU-SILC απαιτούν ένα δείγμα νοικοκυριών αλλά και ένα δείγμα ατόμων ώστε να μελετηθούν όλες. Ας δούμε τις τιμές των ελαχίστων αποτελεσματικών μεγεθών δείγματος για δύο ξεχωριστές περιπτώσεις εφαρμογής της έρευνας. Οι ελάχιστες τιμές αποτελεσματικού μεγέθους δείγματος αφορούν το τμήμα της EU-SILC που αφορά τη μελέτη των σύγχρονων μεταβλητών.

Πίνακας 2: Ελάχιστες τιμές αποτελεσματικού μεγέθους δείγματος στο σύγχρονο τμήμα της έρευνας

	Α.Νοικοκυριά/Διευθύνσεις		Β. Άτομα	
	Α.1.Νοικοκυριά	Α.2.Ερωτώμενα Άτομα ⁷	Β.1.Άτομα-Νοικοκυριά	Β.2.Άτομα άνω των 16 ετών
Βέλγιο	4.750	8.750	6.500	12.000
Βουλγαρία	4.500	10.000		
Τσεχική Δημοκρατία	4.750	10.000	7.500	15.750
Δανία	4.250	7.250	5.500	9.500
Γερμανία	8.250	14.500	11.000	19.250
Εσθονία	3.500	7.750	5.750	12.750
Ελλάδα	4.750	10.000	7.500	15.750
Ισπανία	6.500	16.000	12.000	29.500
Γαλλία	7.250	13.500	10.250	19.000
Ιρλανδία	3.750	8.000	6.000	12.750
Ιταλία	7.250	15.500	11.750	25.000
Κύπρος	3.250	7.500	5.750	13.250
Λετονία	3.750	7.650	6.500	14.750
Λιθουανία	4.000	9.000	6.750	15.250
Λουξεμβούργο	3.250	6.500	5.000	10.000
Ουγγαρία	4.750	10.250	7.750	16.750
Μάλτα	3.000	7.000	5.250	12.250

⁷ Ηλικίας άνω των 16 ετών

Ολλανδία	5.000	8.750	6.500	11.500
Αυστρία	4.500	8.750	6.500	12.750
Πολωνία	6.000	15.000	11.250	28.250
Πορτογαλία	4.500	10.500	8.000	18.750
Ρουμανία	5.250	12.750		
Σλοβενία	3.750	9.000	6.750	16.250
Σλοβακία	4.250	11.000	8.250	21.250
Φινλανδία	4.000	6.750	5.000	8.500
Σουηδία	4.500	7.500	5.750	9.500
Ηνωμένο Βασίλειο	7.500	13.750	10.250	18.750
Σύνολο ΕΕ κράτη μέλη	130.750	272.900		
Ισλανδία	2.250	3.750	3.000	5.000
Νορβηγία	3.750	6.250	4.750	8.000
Σύνολο	136.750	282.900		

Αξίζει να σημειωθούν τα εξής:

- ✚ Οι τιμές των δύο πρώτων στηλών αφορούν πληροφορίες που καταγράφηκαν το 2011.
- ✚ Όσον αφορά τις τιμές της στήλης Β που αφορούν χώρες που χρησιμοποιούν δείγμα ατόμων μόνο οι τιμές της Δανίας, Ολλανδίας, Σλοβενίας, Φινλανδίας, Σουηδίας, Ισλανδίας και Νορβηγίας βρέθηκαν καταγεγραμμένες για το 2011. Οι υπόλοιπες αφορούν τιμές του 2001.
- ✚ Από το 2001 έως και την ημερομηνία καταγραφής του παραπάνω πίνακα μόνο για την περίπτωση της Λετονίας για τη στήλη Α2 η τιμή άλλαξε από 8.500 σε 7.650 το 2011.

Ο υπολογισμός των ελάχιστων αποτελεσματικών δειγμάτων έγινε βάσει των απαιτήσεων που δημοσιεύτηκαν στην Επίσημη Εφημερίδα της Ευρωπαϊκής Ένωσης τον Ιούλιο του 2003, των τεχνικών περιγραφών που δημοσίευσε η Eurostat και του Άρθρου EU-SILC Sampling Guidelines (Verma, 2001).

Η μελέτη των μεταβλητών που αφορούν το εισόδημα στην EU-SILC μελετώνται για κάθε μέλος του νοικοκυριού που έχει επιλεγθεί ενώ η μελέτη των «κοινωνικών μεταβλητών» γίνεται εναλλακτικά και για ένα δείγμα ατόμων (π.χ. ένα μέλος ανά νοικοκυριό).

Ο πίνακας δίνει το ελάχιστο αποτελεσματικό μέγεθος του δείγματος νοικοκυριών (στήλη Α.1.) που απαιτείται έτσι ώστε τόσο οι μεταβλητές του εισοδήματος όσο και οι «κοινωνικές μεταβλητές» να μπορούν να μελετηθούν εφόσον όλα τα μέλη των νοικοκυριών του δείγματος ανταποκριθούν στην έρευνα. Δίνεται επίσης και ο αναμενόμενος αριθμός ατόμων που θα πρέπει να ανταποκριθούν στην έρευνα.(στήλη Α.2.) Ο αριθμός αυτός υπολογίζεται σύμφωνα με το μέγεθος των νοικοκυριών της στήλης Α.1. όπου το μέγεθος των νοικοκυριών αποτελούν τα άτομα ηλικίας άνω των 16 χρόνων

Η στήλη Β του πίνακα αφορά τις χώρες που έχουν επιλέξει ένα δείγμα ατόμων για τη μελέτη των «κοινωνικών μεταβλητών» της έρευνας. Στη στήλη Β.1 δίνεται το ελάχιστο αποτελεσματικό μέγεθος ενός δείγματος ατόμων για τη μελέτη των «κοινωνικών μεταβλητών» και ως εκ τούτου τον αριθμό των νοικοκυριών αφού ένα μέλος ερωτάται από κάθε νοικοκυριό. Στις χώρες αυτές οι πληροφορίες για το εισόδημα προκύπτουν από τα μητρώα των χωρών αφού για τα νοικοκυριά που ανήκουν στο δείγμα εντοπίζεται ο αριθμός των μελών του νοικοκυριού που έχουν ηλικία άνω των 16 ετών. Έτσι δίνεται στον πίνακα ο εκτιμημένος αριθμός αυτών των ατόμων, δηλαδή το μέγεθος των νοικοκυριών σε πλήρη αντιστοιχία με τον υπολογισμό της στήλης Α.2.(στήλη Β.2.)

Μέγεθος υπολογισμού του αληθινού μεγέθους δείγματος

Το ελάχιστο αποτελεσματικό μέγεθος ενός δείγματος είναι το μέγεθος που απαιτείται αν η έρευνα βασίζεται σε απλή τυχαία δειγματοληψία. Τότε, ορίζουμε την επίδραση του σχεδιασμού της έρευνας να παίρνει τη τιμή 1.0. Σε περίπτωση τώρα, που η έρευνα δε βασίζεται σε απλή τυχαία δειγματοληψία (γεγονός που ισχύει στην EU-SILC) το αληθινό μέγεθος του δείγματος

- θα εξαρτάται από το σχεδιασμό του δείγματος (έτσι ώστε η επίδραση να μην υπερβαίνει το 1.0),
- θα πρέπει να καταπολεμά τη περίπτωση της μη ανταπόκρισης των ερωτηθέντων,
- και να ανταποκρίνεται στις εκάστοτε απαιτήσεις των χωρών.

Κάθε χώρα που εφαρμόζει την έρευνα οφείλει να προσαρμόσει το αληθινό μέγεθος του δείγματός σύμφωνα με τις δικές της ανάγκες. Η επίδραση του σχεδιασμού εξαρτάται από το σχεδιασμό του δείγματος και συνεπώς από τις ανάγκες κάθε χώρας που εφαρμόζει την έρευνα. Οι χώρες συνεπώς, αφού επιλέξουν το σχεδιασμό του δείγματος και στη συνέχεια υπολογίσουν την επίδραση του σχεδιασμού στον υπολογισμό των σημαντικότερων μεταβλητών που θέλουν να μελετήσουν καθορίζουν τη τιμή του αληθινού δείγματος σε σχέση με την επίδραση που υπολόγισαν προηγούμενων και το ελάχιστο αποτελεσματικό μέγεθος που έχει προταθεί. Η σχέση αυτή ορίζεται ως εξής:

Έστω n_e το ελάχιστο αποτελεσματικό μέγεθος δείγματος, n_a το αληθινό μέγεθος του δείγματος και $deft$ η επίδραση του σχεδιασμού (από τα αρχικά της αγγλικής ορολογίας - design effect). Τότε:

$$n_e = \frac{n_a}{deft^2}.$$

Σύμφωνα με το τύπο αυτό το μέγεθος του αληθινού δείγματος αυξάνεται όσο αυξάνεται ο βαθμός επίδρασης του σχεδιασμού ενώ στην περίπτωση της τυχαίας δειγματοληψίας όπου ο βαθμός σχεδιασμού παίρνει τη τιμή 1.0 τότε $n_e = n_a$.

Φυσικά το αληθινό μέγεθος θα πρέπει να είναι ακόμη μεγαλύτερο έτσι ώστε να συμπεριλάβει εκ των προτέρων τη μη ανταπόκριση των ερωτηθέντων. Πιο συγκεκριμένα:

Έστω n_s το μέγεθος που θα επιλεγεί αφού έχει υπολογιστεί το αληθινό μέγεθος του δείγματος n_a και R το συνολικό ποσοστό ανταπόκρισης, Τότε:

$$n_s = \frac{n_a}{R}.$$

Ακόμη το τελικό μέγεθος μπορεί να αυξηθεί περαιτέρω έτσι ώστε να καλύψει για παράδειγμα την ανάγκη μίας χώρας για πολύ λεπτομερή περιφερειακή κατανομή. Τελικά το μέγεθος του δείγματος καταγράφεται στην αναφορά ποιότητας που παραδίδει η κάθε χώρα ενώ συνήθως επιλέγεται ένα σαφώς μεγαλύτερο μέγεθος από αυτό του ελαχίστου αποτελεσματικού μεγέθους.

2.2.2. Η λογική επιλογής του ελαχίστου αποτελεσματικού μεγέθους δείγματος

2.2.2.1. Το τμήμα της EU-SILC που μελετά τις σύγχρονες μεταβλητές

Το τελικό μέγεθος του ελάχιστου αποτελεσματικού μεγέθους εκτιμήθηκε μέσα από πολλούς παράγοντες. Αναφέρονται οι εξής:

1. Η εμπειρία της EHCP έρευνας.
2. Ο βαθμός ακρίβειας της εκτίμησης του ποσοστού φτώχειας.
3. Ο βαθμός ακρίβειας της εκτίμησης του μέσου εισοδήματος ανά νοικοκυριό.
4. Το προκαθορισμένο συνολικό μέγεθος των δειγμάτων.
5. Η κατανομή του συνολικού μεγέθους στις χώρες.

1. Μία πρώτη εικόνα για το ελάχιστο αποτελεσματικό μέγεθος υπήρχε λόγω της προηγούμενης έρευνας (EHCP) την οποία αντικατέστησε η έρευνα που μελετάμε. Η ECHP (European Community Household Panel) είχε στόχο να μελετήσει ένα μεγάλο εύρος ζητημάτων που αφορούσαν τις συνθήκες διαβίωσης. Οι πληροφορίες συγκεντρώνονταν από ένα δείγμα νοικοκυριών και ατόμων. Από την έρευνα αυτή η οποία διήρκησε 8 χρόνια φάνηκε ότι ένα δείγμα νοικοκυριών της τάξεως των 4.000 – 6.000 ήταν αρκετό για να οδηγήσει σε πολύ χρήσιμα αποτελέσματα. Το μέγεθος αυτό ήταν επαρκές για να παρέχει ικανοποιητικές εκτιμήσεις για μεταβλητές που αφορούσαν τις συνθήκες ζωής, το εισόδημα, το ποσοστό φτώχειας και τον κοινωνικό αποκλεισμό.

2. Παρόμοιο μέγεθος δείγματος νοικοκυριών εκτιμήθηκε από την εξέταση των απαιτήσεων ακρίβειας της μεταβλητής που αφορά το ποσοστό της φτώχειας. Στην έρευνα ECHP το ποσοστό φτώχειας, που υπολογίστηκε ως το ποσοστό των νοικοκυριών που έχουν εισόδημα 60% μικρότερο από τη διάμεσο της τιμής του εισοδήματος, υπολογίστηκε να κυμαίνεται περίπου από 5%-25%. Διάμεσος είναι η τιμή για την οποία το 50% του πληθυσμού έχει μεγαλύτερη τιμή εισοδήματος από τη τιμή αυτή. Παίρνοντας ως δεδομένο ότι η τιμή του ποσοστού είναι 15% τότε υπολογίστηκε ότι απαιτείται ένα δείγμα μεγέθους 5.000 νοικοκυριών (τυχαία δειγματοληψία) για να εκτιμήσει τη τιμή αυτή με σφάλμα 1%. Είναι σημαντικό να τονίσουμε ότι το μέγεθος του δείγματος δεν είναι ευαίσθητο στην συγκεκριμένη τιμή

του ποσοστού φτώχειας. Αντίστοιχες μετρήσεις για άλλη τιμή του ποσοστού φτώχειας με τον ίδιο αριθμό δείγματος εκτίμησαν τη τιμή στο βαθμό ακρίβειας που επιζητούνταν. Η τάξη μεγέθους του τυπικού σφάλματος ήταν κοινή αλλά διέφεραν οι τιμές.

Πίνακας 3: Εκτίμηση του ποσοστού φτώχειας από δείγμα 5000 νοικοκυριών

Αποτελεσματικό μέγεθος δείγματος 5000 νοικοκυριών		
Ποσοστό φτώχειας (%)	Τυπικό σφάλμα (%)	95% Διάστημα Εμπιστοσύνης
5,0	0,31	4,4-5,6
10,0	0,42	9,2-10,8
15,0	0,50	14,0-16,0
20,0	0,57	18,9-21,1
25,0	0,61	23,8-26,2

3. Εξίσου σημαντική μεταβλητή με αυτή του ποσοστού φτώχειας είναι και η μεταβλητή που αφορά τη μέση τιμή του εισοδήματος των νοικοκυριών. Συνεπώς, μελετήθηκε και η εφαρμογή του μεγέθους του δείγματος για τη μεταβλητή του μέσου εισοδήματος νοικοκυριών. Υπολογίστηκε ότι αν ο συντελεστής διακύμανσης ισούται με 0,7 τότε ένα δείγμα νοικοκυριών τάξεως μεγέθους 5.000 δίνει σχετικό σφάλμα της τάξεως του 1% στο εκτιμημένο μέσο εισόδημα των νοικοκυριών. Ο συντελεστής μεταβλητότητας (coefficient of variation-cv) είναι ο λόγος του τυπικού σφάλματος (σ) προς την απόλυτη μέση τιμή του εισοδήματος (μ). Δηλαδή: $cv = \frac{\sigma}{|\mu|}$.

Το σχετικό σφάλμα του συντελεστή μεταβλητότητας (r.e.) υπολογίζεται ως εξής:

$$r.e. = \frac{cv}{\sqrt{n}}, \text{ n=μέγεθος δείγματος}$$

Συνεπώς η τάξη μεγέθους του σχετικού σφάλματος είναι ικανοποιητική για ένα μέγεθος δείγματος των 5.000 νοικοκυριών.

4. Το ελάχιστο μέγεθος του δείγματος των νοικοκυριών έπρεπε να υπολογιστεί λαμβάνοντας υπόψη ότι το σύνολο των μεγεθών των δειγμάτων έπρεπε να είναι προκαθορισμένο. Συγκεκριμένα, το 2001 δημοσιεύθηκε στην Ευρωπαϊκή Εφημερίδα ότι για τις ΕΕ-15 το άθροισμα των μεγεθών των δειγμάτων έπρεπε να είναι 80.000. Με το μέτρο αυτό φαίνεται ότι το μέγεθος των δειγμάτων των χωρών έπρεπε να περιοριστεί ανάλογα με τον πληθυσμό.

5. Το τελευταίο και σημαντικότερο ζήτημα αφορά την κατανομή του συνολικού μεγέθους στις χώρες που εφαρμόζουν την έρευνα. Παρόλο που για διαφορετικές χώρες με διαφορετικές απαιτήσεις και με άνισο πληθυσμό απαιτούνταν ίδιας τάξης μεγέθους δείγμα είναι φανερό ότι για χώρες με μεγαλύτερο πληθυσμό θα πρέπει να αντιστοιχεί μέγεθος δείγματος μεγαλύτερο από αυτό που αντιστοιχεί σε χώρες με

μικρότερο πληθυσμό. Ένας ακόμη λόγος για αύξηση της τάξης μεγέθους του δείγματος όταν αυξάνεται ο πληθυσμός μιας χώρας είναι και η έκθεση των αποτελεσμάτων της έρευνας σε ευρωπαϊκό επίπεδο. Δεν απαιτείται βέβαια ανάλογη αύξηση λόγω αυτής της έκθεσης. Κάτι τέτοιο όμως θα ήταν μη αποδεκτό αφού για την έκθεση των αποτελεσμάτων σε εθνικό επίπεδο απαιτείται περίπου ίδια τάξη μεγέθους των δειγμάτων της κάθε χώρας. Συνεπώς, το μέγεθος του δείγματος εξαρτάται από τη “σχετική σημασία” (relative importance) της έκθεσης σε εθνικό επίπεδο σε σχέση με την έκθεση σε ευρωπαϊκό επίπεδο. Σχετική σημασία ορίζεται η συνεισφορά ενός παράγοντα σε ένα γενικό γραμμικό μοντέλο. Για να οριστεί η σωστή κατανομή των μεγεθών των δειγμάτων κατασκευάστηκε μία σχέση σύμφωνα με την οποία:

Έστω n_i το ελάχιστο αποτελεσματικό μέγεθος δείγματος για την χώρα i , n_0 ορίζεται μία τιμή έτσι ώστε το σύνολο των μεγεθών να είναι το επιζητούμενο, k είναι η παράμετρος η οποία ορίζεται σύμφωνα με τη σχετική σημασία σε εθνικό επίπεδο σε σχέση με τη σχετική σημασία σε ευρωπαϊκό επίπεδο, για τις τιμές που δίνονται στον Πίνακα 2 η τιμή του $k^2 = 0.25$, M_i είναι το σχετικό μέτρο του πληθυσμού της χώρας. Η τιμή του είναι κανονικοποιημένη έτσι ώστε η μέση τιμή του να είναι 1.0. Τότε δίνεται ότι ο υπολογισμός του ελάχιστου μεγέθους δείγματος βασίζεται στην παρακάτω εξίσωση:

$$n_i = n_0 \sqrt{k^2 + (-k^2) M_i}$$

Οι τιμές του Πίνακα 2 βασίζονται στην παραπάνω εξίσωση αλλά με ένα μικρό περιορισμό του M_i . Ακριβέστερα το M_i αντικαθίσταται από το M_i^a όπου η παράμετρος a εισήχθη για να περιορίσει το μέγιστο μέγεθος δείγματος που αντιστοιχεί στη χώρα με το μεγαλύτερο πληθυσμό, $a = 0.75$. Επιπρόσθετα έχουμε ότι $\sum n_i = 130.750$, $n_0 = 5.785$, $n_{\min} = kn_0 = 3.000$.

Η παραπάνω περιγραφή αφορά κυρίως τις περισσότερες περιπτώσεις εφαρμογής της έρευνας όπου οι χώρες που την εφαρμόζουν χρησιμοποιούν ένα δείγμα νοικοκυριών (στήλη A1). Για τις χώρες οι οποίες χρησιμοποιούν μητρώα ή κάποιο άλλο μέσο καταγραφής ατόμων και επιλέγουν ένα δείγμα ατόμων (στήλη B1) για την έρευνά τους το ελάχιστο αποτελεσματικό μέγεθος θα είναι μικρότερο από αυτό της στήλης A2 όπου όλα τα μέλη του νοικοκυριού παίρνουν μέρος στην έρευνα. Το δείγμα επιλογής ατόμων θεωρήθηκε πιο αποτελεσματικό εξαιτίας του ότι αποφεύγεται η ομαδοποίηση των μελών των νοικοκυριών. Όσον αφορά το δείγμα ατόμων της στήλης A2 όλα τα μέλη των νοικοκυριών ερωτώνται με αποτέλεσμα η ομαδοποίηση αυτών να κρίνεται αναγκαία. Η ομαδοποίηση των ατόμων αυξάνει συνεπώς την επίδραση του σχεδιασμού του δείγματος (defit). Ακριβέστερα η εμπειρία έδειξε ότι ένα δείγμα 75% του εκτιμημένου δείγματος της στήλης A2 είναι ένα επαρκές δείγμα ατόμων για τη μελέτη των μεταβλητών της έρευνας. Τα μεγέθη της

στήλης B1 του πίνακα θα είναι σαφώς μεγαλύτερα από αυτά της στήλης A1 διότι στην περίπτωση του δείγματος ατόμων συλλέγονται λίγες πληροφορίες επιπέδου «νοικοκυριών» αφού οι πληροφορίες αυτές συλλέγονται από ένα άτομο του νοικοκυριού. Σε αντίθεση, το μέγεθος των ατόμων ηλικίας 16 ετών και άνω (στήλη B2) τα οποία εξετάζονται σύμφωνα με τα νοικοκυριά που έχουν επιλεγεί από τη στήλη B1 (βάσει δείγματος ατόμων) και από τα οποία λεπτομερείς πληροφορίες μπορούν να συλλεχθούν είναι σαφώς μεγαλύτερο από αυτό της στήλης A2. Αυτό είναι αποδεκτό λόγω του περιορισμένου κόστους που έχει η συλλογή των πληροφοριών αυτών αφού βασίζεται σε καταγεγραμμένες πληροφορίες και όχι σε προσωπική συνέντευξη. Τέλος στην περίπτωση των χωρών που επιλέγουν ένα δείγμα ατόμων (στήλη B1) και στη συνέχεια από το δείγμα αυτό επιλέγουν το δείγμα νοικοκυριών από τα οποία τα μέλη ηλικίας 16 ετών και άνω εξετάζονται (στήλη B2) λόγω του κόστους της προσωπικής συνέντευξης των ατόμων του δείγματος στόχος του επιλεγόμενου δείγματος είναι να είναι όσο αποτελεσματικό γίνεται. Συνεπώς επιζητείτε ένα όσο το δυνατό ίσης πιθανότητας επιλογής δείγμα ατόμων. Άρα ένα δείγμα νοικοκυριών μη ίσης πιθανότητας επιλογής όπως έχει συζητηθεί στην παράγραφο «Επιλογή Δείγματος». Κατά συνέπεια το δείγμα νοικοκυριών είναι λιγότερο αποτελεσματικό σε σχέση με αυτό που επιλέγεται στην περίπτωση της στήλης A1.

Συγκεντρωτικά έχουμε:

- ✓ Μέγεθος στήλης B1 > Μέγεθος στήλης A1.
- ✓ Μέγεθος στήλης B1 < Μέγεθος στήλης A2.
 - ✓ Μέγεθος στήλης B1 = 75% Μεγέθους στήλης A2.
- ✓ Μέγεθος στήλης B2 > Μέγεθος στήλης A2.

Το μέγεθος στήλης αφορά φυσικά το ελάχιστο αποτελεσματικό μέγεθος δείγματος σε κάθε περίπτωση.

2.2.2.2. Το τμήμα της EU-SILC που μελετά τις διαχρονικές μεταβλητές

Στο σημείο αυτό είναι σημαντικό να περιγράψουμε τη λογική σύμφωνα με την οποία καθορίστηκαν τα ελάχιστα αποτελεσματικά δείγματα ώστε να εξυπηρετούν εξίσου το τμήμα της EU-SILC που μελετάει τις διαχρονικές μεταβλητές. Προτεραιότητα της EU-SILC είναι να κατασκευάσει ένα αρκετά μεγάλο και αντιπροσωπευτικό δείγμα που αφορά τη μελέτη των σύγχρονων μεταβλητών. Σύμφωνα με το σχεδιασμό του δείγματος που κυρίως χρησιμοποιείται από τις χώρες που εφαρμόζουν την έρευνα (εναλλασόμενος σχεδιασμός) το διαχρονικό τμήμα συνδέεται με το σύγχρονο τμήμα. Το αποτέλεσμα αυτής της σύνδεσης είναι τα δύο τμήματα της έρευνας που αφορούν το δείγμα να μην μπορούν να καθοριστούν ανεξάρτητα. Το μεγαλύτερο κόστος και οι περιορισμοί του μεγέθους του δείγματος εφαρμόζονται στο διαχρονικό τμήμα της έρευνας. Συνεπώς, η βασική προσέγγιση

ώστε να καθοριστούν τα ελάχιστα αποτελεσματικά μεγέθη δειγμάτων για το διαχρονικό τμήμα της έρευνας είναι η ίδια με αυτή που μόλις περιγράφηκε για το σύγχρονο τμήμα της έρευνας.

Όσον αφορά το βαθμό ακρίβειας που απαιτείται π.χ. για τη σημαντική μεταβλητή αυτού του τμήματος που αφορά το ποσοστό των ανθρώπων που παραμένουν στη φτώχεια (persistent poverty rate), το οποίο καθορίζεται από το ποσοστό των ανθρώπων που παραμένουν στη φτώχεια (δηλαδή έχουν κατά 60% εισόδημα μικρότερο από τη τιμή της διαμέσου του εισοδήματος) για τουλάχιστον ένα αριθμό συνεχόμενων ετών. Σύμφωνα με την ECHP έρευνα ένα ποσοστό 10% των ανθρώπων παραμένουν στη φτώχεια για δύο ή περισσότερα έτη. Λαμβάνοντας υπόψη την ίδια απαίτηση ακρίβειας όπως και για τις σημαντικές μεταβλητές του σύγχρονου τμήματος της έρευνας (σφάλμα της τάξεως του 1%) ένα αποτελεσματικό μέγεθος λίγο μικρότερο από 4.000 νοικοκυριά απαιτείται. Στηριζόμενοι σε αυτή τη βάση δείγματος και στο γεγονός ότι οι παράγοντες που επηρέασαν το τελικό ελάχιστο αποτελεσματικό μέγεθος του σύγχρονου τμήματος της έρευνας είναι κοινά, αποφασίστηκε ότι τα μεγέθη δειγμάτων που αφορούν το διαχρονικό τμήμα της έρευνας είναι το 75% του αντίστοιχου ελάχιστου μεγέθους δείγματος που υπολογίστηκε για το σύγχρονο τμήμα της έρευνας.

Βέβαια, η μεταβλητή του παραμένουσας ποσοστού ανθρώπων στη φτώχεια είναι αρκετά πολύπλοκη και δεν εξαρτάται μόνο από το πόσοι άνθρωποι παρέμειναν να έχουν εισόδημα κατά τουλάχιστον 60% μικρότερο από τη διάμεσο. Το ίδιο ισχύει και για τη μεταβλητή του ποσοστού φτώχειας στο σύγχρονο τμήμα αλλά όχι σε τέτοιο βαθμό ώστε η απλοποίηση του να μη αντικειμενική για το σκοπό που γίνεται. Για την ακρίβεια, η μεταβλητή του ποσοστού παραμένουσας φτώχειας εμφανίζει πολύ μεγαλύτερο εύρος διασποράς σε σχέση με αυτό του ποσοστού φτώχειας. Κατά συνέπεια η μέση τιμή του ποσοστού φτώχειας που χρησιμοποιήθηκε για τη βάση της μέτρησης μπορεί να θεωρηθεί εξαιρετική απλοποίηση μολονότι τα αποτελέσματα δείχνουν ότι δεν είναι ευαίσθητα στο γεγονός στο να μετριάσουν τις διασπορές στο ποσοστό αυτό. Η εφαρμογή όμως έδειξε ότι χρησιμοποιώντας το ελάχιστο αποτελεσματικό δείγμα που καθορίστηκε προηγουμένως, το σχετικό σφάλμα που αφορά την εκτίμηση του ποσοστού παραμονής στη φτώχεια τείνει να είναι μεγαλύτερο στις χώρες οι οποίες έχουν χαμηλά σύγχρονα ποσοστά φτώχειας. Ευτυχώς, ένα μέρος αυτών των χωρών αποτελούν χώρες που διατηρούν μητρώα και ως εκ τούτου το μέγεθος του δείγματος που αφορά τις μεταβλητές του εισοδήματος – ή μεταβλητές που το αφορούν – μπορεί να αυξηθεί εύκολα. (στήλη B2). Έτσι κι αλλιώς υπάρχει μία σχετική ελευθερία όσον αφορά το μέγεθος του δείγματος που αφορά τη στήλη B2 για το σκοπό αυτό. Στον παρακάτω πίνακα μπορούμε να δούμε τις τιμές των ελαχίστων αποτελεσματικών δειγμάτων όπως βρέθηκαν το 2011.

Πίνακας 4: Ελάχιστες τιμές αποτελεσματικού μεγέθους δείγματος στο διαχρονικό τμήμα της έρευνας

	Α.Νοικοκυριά/Διευθύνσεις		Β. Άτομα	
	Α.1.Νοικοκυριά	Α.2.Ερωτώμενα Άτομα ⁸	Β.1.Νοικοκυριά & Επιλεγμένοι ερωτηθέντες	Β.2.Άτομα άνω των 16 ετών
Βέλγιο	3.500	6.500		
Βουλγαρία	3.500	7.500		
Τσεχική Δημοκρατία	3.500	7.500		
Δανία	3.250	5.500	4.250	7.250
Γερμανία	6.000	10.500		
Εσθονία	2.750	5.750		
Ελλάδα	3.500	7.250		
Ισπανία	5.000	12.250		
Γαλλία	5.500	10.250		
Ιρλανδία	2.750	6.000		
Ιταλία	5.500	11.750		
Κύπρος	2.500	5.500		
Λετονία	2.750	5.600		
Λιθουανία	3.000	6.750		
Λουξεμβούργο	2.500	6.000		
Ουγγαρία	3.500	7.750		
Μάλτα	2.250	5.250		
Ολλανδία	3.750	6.500	5.000	8.750
Αυστρία	3.250	6.250		
Πολωνία	4.500	11.250		
Πορτογαλία	3.250	7.500		
Ρουμανία	4.000	9.500		
Σλοβενία	2.750	6.750	5.000	12.250
Σλοβακία	3.250	8.250		
Φινλανδία	3.000	5.000	3.750	6.250
Σουηδία	3.500	5.750	4.500	7.500
Ηνωμένο Βασίλειο	5.750	10.500		
Σύνολο ΕΕ κράτη μέλη	98.250	203.850		
Ισλανδία	1.700	2.800	2.000	3.250
Νορβηγία	2.750	4.650	3.500	6.000
Σύνολο	102.700	211.300		

⁸ Ηλικίας άνω των 16 ετών

Κεφάλαιο 3

Ποιοτική Έκθεση (Quality Report)

Για την παραγωγή εθνικών στατιστικών αποτελεσμάτων και ζητημάτων είναι υπεύθυνες οι εθνικές στατιστικές υπηρεσίες των χωρών οι οποίες ελέγχουν και αναλύουν τα δεδομένα που συλλέγονται πριν τα μεταφέρουν στη Ευρωπαϊκή Στατιστική Υπηρεσία, τη Eurostat. Η Eurostat από την άλλη πλευρά, όπως έχει προαναφερθεί, αναλαμβάνει την επεξεργασία των στατιστικών πληροφοριών και τη δημοσίευση αυτών σε ευρωπαϊκό επίπεδο καθώς επίσης και τη διασφάλιση ότι οι πληροφορίες αυτές είναι συγκρίσιμες, με το να δημιουργεί μία εναρμονισμένη μεθοδολογία για όλες τις χώρες. Είναι φανερό ότι η συνεργασία των εθνικών στατιστικών υπηρεσιών με τη Eurostat είναι αναγκαία.

Η απαιτούμενη συνεργασία μεταξύ των εθνικών στατιστικών υπηρεσιών ή και άλλων εθνικών αρχών με τη Eurostat οδήγησε στη δημιουργία μίας μη επίσημης δομής, χωρίς προηγούμενο στην Ευρωπαϊκή Επιτροπή, το επονομαζόμενο Ευρωπαϊκό Στατιστικό Σύστημα (ESS- European Statistical System). Το Ευρωπαϊκό Στατιστικό Σύστημα κατασκευάστηκε σταδιακά με σκοπό την παροχή αξιόπιστων και συγκρίσιμων στατιστικών πληροφοριών σε ευρωπαϊκό επίπεδο. Το ESS λειτουργεί ως δίκτυο στο οποίο ο ρόλος της Eurostat είναι ο εναρμονισμός των στατιστικών πληροφοριών σε συνεργασία με τις εθνικές στατιστικές υπηρεσίες.

Η Ποιοτική Έκθεση παρουσιάζει πληροφορίες που αφορούν την ποιότητα των στοιχείων που δημοσιεύονται από μία έρευνα καθώς επίσης και πληροφορίες που αφορούν τα εργαλεία και τα πρότυπα μίας ποιοτικής έκθεσης τα οποία έχουν συμφωνηθεί εντός του Ευρωπαϊκού Στατιστικού Συστήματος.

Μία ποιοτική έκθεση μπορεί να καταγράφεται σε εθνικό επίπεδο και να αφορά τα ποιοτικά χαρακτηριστικά της έρευνας που διεξήχθη αλλά και σε ευρωπαϊκό επίπεδο, δηλαδή μία συγκεντρωτική ποιοτική έκθεση όλων των χωρών που έλαβαν μέρος στην έρευνα. Για παράδειγμα, στην περίπτωση της EU-SILC αναφέρθηκε ότι ζητείται ενδιάμεση ποιοτική έκθεση όταν συλλέγονται τόσο τα σύγχρονα όσο και τα διαχρονικά δεδομένα. Βεβαίως, σε μία ποιοτική ανάλυση ευρωπαϊκού επιπέδου, για την ολοκλήρωση της απαιτείται μία ποιοτική έκθεση από κάθε χώρα που έλαβε μέρος στην έρευνα. Σε κάθε περίπτωση μία ποιοτική έκθεση περιλαμβάνει τομείς που αφορούν την ποιότητα και είναι κοινοί για οποιαδήποτε ποιοτική έκθεση συνταχθεί. Η χρησιμότητα και η σημαντικότητα των ποιοτικών εκθέσεων θα γίνει φανερή από την ανάλυση των τομών που περιλαμβάνονται.

Δομή μίας ποιοτικής έκθεσης

Η δομή μίας ποιοτικής έκθεσης είναι αυστηρά καθορισμένη σύμφωνα με τα πρότυπα του Ευρωπαϊκού Στατιστικού Συστήματος με αποτέλεσμα να είναι κοινή οποτεδήποτε συντάσσεται, και η πρόσβαση στις πληροφορίες που περιέχονται να είναι άμεση. Φυσικά η σύνταξη και η μορφή της ποιοτικής έκθεσης επιλέγεται

σύμφωνα με το τύπο και τα χαρακτηριστικά της στατιστικής διαδικασίας που πρόκειται να περιγραφεί.

Στόχοι που καλύπτονται από την αυστηρή δομή μίας ποιοτικής έκθεσης

Όπως προαναφέρθηκε η δομή της ποιοτικής έκθεσης είναι αυστηρά καθορισμένη με αποτέλεσμα να είναι κοινή για οποιονδήποτε σκοπό συντάσσεται. Οι βασικοί στόχοι που καλύπτονται μέσα από τη δομή της έκθεσης είναι:

- Μέσα από την προώθηση εναρμονισμένων ποιοτικών εκθέσεων που αφορούν όλες τις στατιστικές διαδικασίες και τα αποτελέσματα αυτών σε ένα κράτος μέλος της Ευρωπαϊκής Ένωσης (ή μη) διευκολύνεται η επίτευξη σύγκρισης μεταξύ των στατιστικών διαδικασιών αλλά και των αποτελεσμάτων. Άλλωστε η επίτευξη σύγκρισης μεταξύ τόσο των διαδικασιών αλλά και των αποτελεσμάτων είναι πρωταρχικός λόγος που συντάσσονται οι ευρωπαϊκές έρευνες.
- Επιτυγχάνεται η σύγκριση μεταξύ κρατών μελών τα οποία έχουν εφαρμόσει παρόμοιες στατιστικές διαδικασίες κατά τη διάρκεια της έρευνας. Ομοίως επιτυγχάνεται σύγκριση μεταξύ των κρατών μελών που οδηγούνται σε κοινά αποτελέσματα.
- Η αυστηρή δομή της ποιοτικής έκθεσης διασφαλίζει ότι οι εκθέσεις που παραδίδονται θα περιέχουν όλες τις απαραίτητες πληροφορίες για την αναγνώριση των στατιστικών διαδικασιών (Η επεξήγηση του όρου δίνεται στην επόμενη παράγραφο.) αλλά και τυχόν ποιοτικών προβλημάτων που μπορεί να προκύψουν στα αποτελέσματα. Τέλος, η αυστηρή δομή που θα δούμε παρακάτω βοηθά στην αναγνώριση πιθανών βελτιώσεων.

Στο σημείο αυτό είναι αναγκαίο να καταγραφεί η έννοια των στατιστικών διαδικασιών. Με τον όρο *στατιστική διαδικασία* εννοείται μία δειγματοληπτική έρευνα, μία απογραφή, η αξιοποίηση ήδη αποθηκευμένων στοιχείων σε βάσεις δεδομένων (administrative data), ο καθορισμός τιμών ή άλλων οικονομικών δεικτών αλλά και οποιαδήποτε άλλη στατιστική επεξεργασία αναλαμβάνει μία εθνική στατιστική υπηρεσία.

Μία ποιοτική έκθεση περιλαμβάνει ποιοτικές πληροφορίες για την στατιστική διαδικασία που διεξήχθη οι οποίες καταγράφονται στο κατάλληλο τμήμα της έκθεσης. Πιο συγκεκριμένα μία ποιοτική έκθεση περιλαμβάνει τα τμήματα:

1. Εισαγωγή στη στατιστική διαδικασία και στα αποτελέσματά της.
2. Συνάφεια (Relevance).
3. Ακρίβεια (Accuracy).
4. Επικαιρότητα και Εγκαιρότητα (Timeliness and Punctuality).
5. Προσβασιμότητα και Σαφήνεια (Accessibility and Clarity).
6. Συνοχή και Συγκρισιμότητα (Coherence and Comparability).
7. Συμβιβασμοί μεταξύ των ποιοτικών συνιστωσών των αποτελεσμάτων.
8. Εκτίμηση των αναγκών και των αντιλήψεων των χρηστών.
9. Επιδόσεις, Κόστος και Επιβάρυνση Ανταπόκρισης.

10. Εμπιστευτικότητα, Διαφάνεια και Ασφάλεια.

11. Συμπεράσματα.

3.1. Εισαγωγή στη στατιστική διαδικασία και τα αποτελέσματά της

Σκοπός του τμήματος της εισαγωγής σε μία ποιοτική έκθεση είναι η παρουσίαση της στατιστικής διαδικασίας που πρόκειται να αναλυθεί ποιοτικά καθώς και των αποτελεσμάτων της. Στην εισαγωγή περιέχονται οι απαραίτητες πληροφορίες για την αντίληψη της αφού αποτελεί και το θέμα της ποιοτικής έκθεσης. Η παρουσίαση της στατιστικής διαδικασίας ξεκινά με μία *ιστορική αναφορά* της ημερομηνίας που εισήχθη και δρομολογήθηκε η έρευνα αλλά και των αρχικών στόχων της. Στη συνέχεια γίνεται μία συνολική περιγραφή της διαδικασίας, των αποτελεσμάτων της αλλά και της εξέλιξης της στο χρόνο. Έτσι καταγράφονται οι γενικές πληροφορίες οι οποίες είναι απαραίτητες για την αντίληψη του περιεχομένου της στατιστικής διαδικασίας αλλά και των κυρίως αποτελεσμάτων της που πρόκειται να αναλυθούν.

Είναι απαραίτητο να αναφερθεί η *στατιστική πηγή (ή πηγές)* όπου περιέχονται τα αποτελέσματα της στατιστικής διαδικασίας που αναλύεται αλλά και η αναφορά των τυχόν αποτελεσμάτων που ανήκουν στην ίδια πηγή και συγχέονται με τα αποτελέσματα της στατιστικής διαδικασίας που αναφερόμαστε.

Επίσης κρίνεται αναγκαία η *αναφορά των ορίων* της στατιστικής διαδικασίας με τα αποτελέσματα που αναφέρονται στη συγκεκριμένη ποιοτική έκθεση αλλά και των αποτελεσμάτων που δίνονται σε σχετικές ποιοτικές εκθέσεις. Αυτό είναι σημαντικό διότι αρκετές φορές το όριο δεν είναι προφανές. Για παράδειγμα, αρκετές φορές αποτελέσματα με διαφορετικά ονόματα και εννοιολογικούς στόχους μπορεί να έχουν μία ή και περισσότερες κοινές μεταξύ τους υπο-διαδικασίες ή ακόμη να μοιράζονται την ίδια βάση αρχικών δεδομένων (micro database). Στην περίπτωση αυτή οι λόγοι για τους οποίους επιλέχθηκε το συγκεκριμένο όριο πρέπει να δικαιολογηθούν.

Επιπρόσθετα, στο τμήμα της εισαγωγής της ποιοτικής έκθεσης θεωρείται αναγκαία μία *συνολική επισκόπηση* των αποτελεσμάτων που συνδέονται με τη στατιστική διαδικασία η οποία θα περιέχει αναφορές των ΜΜΕ, εθνικά αποτελέσματα καθώς και αποτελέσματα που δόθηκαν σε διεθνείς οργανισμούς αλλά και τα αποτελέσματα που δόθηκαν στο Ευρωπαϊκό Στατιστικό Σύστημα.

Ακόμη θεωρήθηκε σημαντική η *αναφορά των παραπομπών* (κυρίως με τη χρήση υπερσυνδέσμων) σε άλλα έγγραφα κυρίως όταν οι παραπομπές αφορούν τη μεθοδολογία της στατιστικής διαδικασίας ή την ποιότητα των αποτελεσμάτων.

Τέλος, κρίνεται σημαντικό να υποδεικνύονται στο τμήμα αυτό τα σημαντικότερα ποιοτικά προβλήματα που εμφανίστηκαν στη στατιστική διαδικασία.

Είναι φανερό ότι το τμήμα της εισαγωγής στη ποιοτική έκθεση είναι αναγκαίο για το χρήστη της έκθεσης. Συνολικά, έχουμε ότι στο τμήμα της εισαγωγής στη ποιοτική έκθεση περιλαμβάνονται:

1. Εισαγωγή

- ✓ Ιστορική Αναφορά της στατιστικής διαδικασίας και των αποτελεσμάτων της.
- ✓ Η στατιστική πηγή στην οποία περιέχονται τα αποτελέσματα και τα συνδεδεμένα αποτελέσματα.
- ✓ Τα όρια της στατιστικής διαδικασίας και αναφορές σε άλλες ποιοτικές εκθέσεις.
- ✓ Συνολική Επισκόπηση των αποτελεσμάτων της διαδικασίας.
- ✓ Αναφορές σε άλλα έγγραφα που αφορούν κυρίως το τμήμα της μεθοδολογίας.

3.2. Συνάφεια (Relevance)

Το τμήμα της συνάφειας είναι ένα από τα σημαντικότερα τμήματα μίας ποιοτικής έκθεσης. Ας ξεκινήσουμε με τη σημασία του όρου. *Συνάφεια* είναι ο βαθμός στον οποίο τα αποτελέσματα της στατιστικής διαδικασίας ικανοποιούν τις τρέχουσες και δυνητικές ανάγκες των χρηστών. Η συνάφεια εξαρτάται από το κατά πόσο παρήχθησαν όλες οι στατιστικές (=οτιδήποτε μπορεί να προκύψει μέσα από οποιαδήποτε στατιστική διαδικασία) που χρειαζόντουσαν αλλά και από το βαθμό στον οποίο το σκεπτικό που χρησιμοποιήθηκε (επεξηγήσεις, ομαδοποιήσεις, κ.λπ.) ανταποκρίνεται στις ανάγκες των χρηστών.

Συνεπώς το τμήμα αυτό περιέχει ποιοτικές πληροφορίες που αφορούν τη σύνδεση της στατιστικής διαδικασίας και των αποτελεσμάτων της με τις ανάγκες των χρηστών. Πιο συγκεκριμένα στο τμήμα αυτό ανήκει καταρχάς μία *περιγραφή που βασίζεται στο περιεχόμενο* των αποτελεσμάτων της στατιστικής διαδικασίας που αναλύεται. Η περιγραφή αυτή προτείνεται να περιέχει:

- Βασικούς δείκτες (κυρίως εκείνους που τονίζονται από το τύπο π.χ. εθνική ανεργία, πληθωρισμός κ.α.). Οι βασικοί δείκτες είναι οι δείκτες που στοχεύεται να μελετηθούν μέσα από μία στατιστική διαδικασία.
- Μεταβλητές (π.χ. κύκλος εργασιών, κατανάλωση, εργασία, μισθοί). Οι μεταβλητές χρησιμοποιούνται για τη μελέτη των δεικτών. Οι τιμές τους συγχέονται με τις τιμές των δεικτών που στοχεύετε να μελετηθούν.
- Υποτομείς (sub-domains) / για τους οποίους οι δείκτες εμφανίζονται ξεχωριστά.
- Εκτιμήσεις των επιπέδων των τιμών των αποτελεσμάτων σε σχέση με εκάστοτε αλλαγές (χρονοσειρά).
- Αναφερόμενη περίοδος και συχνότητα παραγωγής στατιστικών

Επίσης, στο τμήμα της συνάφειας κρίνεται σημαντικό να γίνει μία αξιολόγηση των βασικών αποτελεσμάτων / δεικτών που επιθυμούνταν από τους χρήστες καθώς επίσης και μία αναφορά σε οποιαδήποτε ανεπάρκεια προέκυψε στα αποτελέσματα και αφορά σημαντικούς χρήστες αυτών των αποτελεσμάτων. Για

παράδειγμα, το τελευταίο μπορεί να αφορά μη ικανοποιητική υποδιαίρεση των δεδομένων σε υποτομείς, χρονοσειρές που μπορεί να είναι πολύ μικρές σε διάρκεια ή και αποτελέσματα που δεν είναι συχνά, π.χ. τριμηνιαία αντί για μηνιαία. Τα παραπάνω παραδείγματα σαφώς και αποτελούν ελλείψεις για τους χρήστες των αποτελεσμάτων. Είναι φανερό ότι σε μία έρευνα είναι πιθανό να μην ικανοποιηθούν όλες οι ανάγκες των χρηστών των αποτελεσμάτων. Οι λόγοι μπορεί να είναι είτε τεχνικοί είτε οικονομικοί. Παρ'όλα αυτά στο τμήμα της συνάφειας πρέπει να δίνονται πληροφορίες για τη *μη ικανοποίηση όλων των αναγκών* των χρηστών, οι λόγοι για τους οποίους συγκεκριμένες ανάγκες δεν ικανοποιήθηκαν μέσα από τη στατιστική διαδικασία αλλά και τα πλάνα που υπάρχουν για ικανοποίηση αυτών των αναγκών στο μέλλον.

Ακόμη στο τμήμα αυτό πρέπει να δοθούν οι ορισμοί των στατιστικών εννοιών που χρησιμοποιούνται αλλά και η σχέση τους με τους ορισμούς που ιδανικά θα δίνονταν από τους χρήστες. Οι έννοιες που ορίζονται κατά τη διάρκεια του σχεδιασμού και του προγραμματισμού μίας έρευνας αφορούν τον πληθυσμό στον οποίο θα αναφέρεται η έρευνα, την επεξήγηση των μονάδων του πληθυσμού αλλά και τη μέθοδο συνάθροισης των δεδομένων που χρησιμοποιείτε. Συχνά, ο *ορισμός των εννοιών* με αυτό που θεωρείται ιδανικό από τους χρήστες να ορίζεται, διαφέρει. Στην περίπτωση αυτή είναι φανερό ότι στο τμήμα της συνάφειας πρέπει να αναφερθούν αυτές οι διαφορές. Βέβαια ακόμη συχνότερα, το παραπάνω δεν είναι εφικτό, με αποτέλεσμα να επιλέγεται ένας ενιαίος ορισμός, το κίνητρο για την επιλογή αυτού πρέπει να δίνεται στο τμήμα αυτό. Φυσικά, οποιεσδήποτε αποκλίσεις από τους χρησιμοποιούμενους που είναι και αποδεκτοί από το ESS ορισμούς ή από ορισμούς που χρησιμοποιούνται διεθνώς είναι αναγκαίο να αναφέρονται στο τμήμα της συνάφειας μιας ποιοτικής έκθεσης.

Προτείνεται ακόμη, οποτεδήποτε είναι εφικτή ή αριθμητική απεικόνιση των ευαισθησιών των αποτελεσμάτων στον ορισμό των εννοιών, να παρουσιάζεται διότι μπορεί να δώσει σημαντικές πληροφορίες. Η βάση αυτών των απεικονίσεων μπορεί να είναι μία ανάλυση ή προσομοίωση της ευαισθησίας των αποτελεσμάτων. Οι απεικονίσεις αυτές δίνουν στους χρήστες πληροφορίες για τους κινδύνους ενός προβλήματος σχετικότητας που αφορά την ιδιαίτερη αίτησή τους. Με την έννοια αυτή ο χρήστης μπορεί να δει για παράδειγμα τα προβλήματα που θα προκύψουν στα αποτελέσματα αν υπάρχει διαφορά μεταξύ του ορισμού που δόθηκε και του ορισμού που ήθελε ο χρήστης να δοθεί.

Στην περίπτωση που υπάρχουν δείκτες, μεταβλητές που δε μελετήθηκαν επαρκώς ή και καθόλου σύμφωνα με το Ευρωπαϊκό Στατιστικό Σύστημα ή σύμφωνα με τους διεθνείς κανονισμούς/οδηγίες τότε τα στατιστικά στοιχεία που δίνονται θεωρούνται ανεπαρκή. Είναι αναγκαίο λοιπόν να περιέχεται στο τμήμα της συνάφειας μία σαφής δήλωση του βαθμού της ανεπάρκειας, εφ' όσον υπάρχει, ο καθορισμός του οποίου γίνεται σύμφωνα με τους κανονισμούς του ESS. Επιπλέον σημαντικό είναι να παρουσιάζονται και τα πλάνα που υπάρχουν και αφορούν τη βελτίωση του βαθμού αυτού στο μέλλον. Εναλλακτικός τρόπος υπολογισμού της πληρότητας των στατιστικών είναι η σύγκρισή τους με τους εθνικούς στόχους.

Τέλος, είναι σημαντικό στο τμήμα της συνάφειας της ποιοτικής ανάλυσης να δίνονται οι ποιοτικοί δείκτες εφόσον μπορούν να υπολογιστούν. Ποιοτικός δείκτης θεωρείται το ποσοστό των διαθέσιμων στατιστικών. Με την έννοια *στατιστικό* θα εννοείται στο εξής οποιαδήποτε πληροφορία δόθηκε μέσα από μία στατιστική μελέτη. Συνεπώς, ποιοτικός δείκτης μπορεί να είναι για παράδειγμα ο αριθμός των στατιστικών στοιχείων που παραχωρήθηκαν σύμφωνα με ένα σχετικό κανονισμό του ESS σε σχέση με τον αριθμό των στατιστικών αποτελεσμάτων που καθορίζονται από τον κανονισμό. Ο παραπάνω ποιοτικός δείκτης είναι διαθέσιμος σε περίπτωση που υπάρχει ένας ESS κανονισμός ή οδηγός. Στην περίπτωση τώρα που τα στατιστικά στοιχεία δεν είναι ίσης βαρύτητας μεταξύ τους απαιτείται η μέθοδος της στάθμισης, προηγουμένως (η μέθοδος αναφέρθηκε στο τμήμα των απαιτήσεων της EU-SILC, στο 1^ο κεφάλαιο).

Συγκεντρωτικά, έχουμε ότι στο τμήμα της συνάφειας μίας ποιοτικής έκθεσης πρέπει να αναφέρονται πληροφορίες που αφορούν:

2. Συνάφεια

- ✓ Βάσει περιεχομένου περιγραφή των στατιστικών αποτελεσμάτων.
- ✓ Ανάγκες των χρηστών που δεν ικανοποιήθηκαν, και οι λόγοι για τη μη ικανοποίησή τους.
- ✓ Ορισμοί των στατιστικών εννοιών.
- ✓ Πληροφορίες για την πληρότητα συγκρινόμενες με τους σχετικούς κανονισμούς/οδηγίες.
- ✓ Διαθέσιμοι ποιοτικοί δείκτες.

Όπως έχει προαναφερθεί, οι στατιστικές διαδικασίες ποικίλουν με αποτέλεσμα τα ιδιαίτερα χαρακτηριστικά τους να οδηγούν σε επιπλέον καταγραφή στοιχείων στο τμήμα της συνάφειας ανάλογα με τις απαιτήσεις της διαδικασίας.⁹ Τα στοιχεία που προαναφέρθηκαν αποτελούν ένα γενικό πλαίσιο των περιεχομένων που κρίνεται αναγκαίο να περιγράφονται στο τμήμα της συνάφειας μίας ποιοτικής έκθεσης.

3.3. Ακρίβεια (Accuracy)

3.3.1. Γενικό πλαίσιο

Είναι γνωστό ότι ένας από τους σκοπούς των στατιστικών μελετών είναι η εκτίμηση άγνωστων τιμών οι οποίες αφορούν ποσοτικά χαρακτηριστικά ενός πληθυσμού. Το φαινόμενο της *διασποράς* εμφανίζεται όταν οι τιμές που εκτιμήθηκαν από διαφορετικές εφαρμογές της στατιστικής μελέτης απέχουν μεταξύ τους. Το φαινόμενο αυτό οφείλεται σε τυχαίους παράγοντες. Από την άλλη πλευρά το φαινόμενο της μεροληψίας εμφανίζεται όταν ο μέσος όρος της εκτιμήτριας που

⁹ Τα επιπλέον στοιχεία δίνονται στις επίσημες οδηγίες που έχουν δοθεί για τη δομή μίας ποιοτικής έκθεσης.

χρησιμοποιούμε δεν έχει μέση τιμή ίση με τη ποσότητα που εκτιμά. Το φαινόμενο αυτό οφείλεται σε παράγοντες που εμφανίζονται είτε συστηματικά ή όχι. Η μεροληψία μίας εκτιμήτριας συνάρτησης ισούται με τη διαφορά της αναμενόμενης εκτιμημένης τιμής από την αληθινή.

Η *ακρίβεια* των αποτελεσμάτων μίας στατιστικής διαδικασίας είναι ο βαθμός στον οποίο τα αποτελέσματα προσεγγίζουν τις αληθινές τιμές. Ο ορισμός είναι σύμφωνος με τη γενική στατιστική αίσθηση που κυριαρχεί. Συνεπώς, στο τμήμα της ακρίβειας καταγράφονται πληροφορίες που αφορούν και επηρεάζουν το βαθμό της ακρίβειας των αποτελεσμάτων.

Σημαντικό τμήμα των πληροφοριών αυτών αποτελεί ο εντοπισμός των βασικών πηγών *σφαλμάτων* των μεταβλητών. Κατά τη διάρκεια εφαρμογής μίας στατιστικής διαδικασίας εμφανίζονται σφάλματα τα οποία επηρεάζουν την ακρίβεια των αποτελεσμάτων. Τα σφάλματα αυτά είναι ποικίλα με αποτέλεσμα να έχει δημιουργηθεί μία ταξινόμηση αυτών ανάλογα με το περιεχόμενό τους και τη διαδικασία στην οποία εμφανίστηκαν. Πιο συγκεκριμένα υπάρχουν δύο μεγάλες κατηγορίες λαθών τα *δειγματοληπτικά σφάλματα* και τα *μη δειγματοληπτικά σφάλματα*. Τα σφάλματα δειγματοληψίας εμφανίζονται μόνο στις έρευνες που βασίζονται σε δείγμα του πληθυσμού και οφείλονται στο γεγονός ότι επιλέγεται ένα «υποσύνολο» του πληθυσμού με τυχαίο τρόπο. Τα μη δειγματοληπτικά σφάλματα εμφανίζονται σε όλες τις στατιστικές διαδικασίες και χωρίζονται σε υποκατηγορίες. Συνεπώς έχουμε:

- Δειγματοληπτικά Σφάλματα (Sampling Errors).
- Μη δειγματοληπτικά Σφάλματα (Non-Sampling Errors).
 - Σφάλματα Κάλυψης (Coverage Errors).
 - Σφάλματα Μέτρησης (Measurement Errors).
 - Σφάλματα μη Απόκρισης (Non-Response Errors).
 - Σφάλματα Επεξεργασίας (Processing Errors).

Τα *σφάλματα κάλυψης* προκύπτουν από την αποτυχία επαρκούς κάλυψης όλων των κατηγοριών που συναποτελούν ένα πληθυσμό. Στο φαινόμενο αυτό οδηγούν τα ατελή (μη πλήρη) πλαίσια του δείγματος. Πλαίσιο ενός δείγματος θεωρείται η λίστα στην οποία περιέχονται όλα τα μέλη του πληθυσμού και από την οποία γίνεται η δειγματοληψία. Τα σφάλματα κάλυψης εμφανίζονται ουσιαστικά λόγω των διαφοροποιήσεων μεταξύ του πληθυσμού που σκοπεύετε να μελετηθεί σε μία έρευνα και του πλαισίου του δείγματος. Τα σφάλματα κάλυψης μπορεί να είναι α) η υπερκάλυψη, β) η ανεπαρκής κάλυψη και γ) η εσφαλμένη ταξινόμηση. Σύμφωνα με τη περιγραφή των σφαλμάτων κάλυψης η *υπερκάλυψη* προκαλείται από τη εσφαλμένη συμπερίληψη μονάδων του πληθυσμού στη διεξαγωγή της έρευνας. Το γεγονός αυτό μπορεί να οφείλεται στην ύπαρξη ελαττωματικού πλαισίου, στη συμπερίληψη μη προσδιορισμένων μονάδων, ή στη συμπερίληψη προσδιορισμένων μονάδων στη διεξαγωγή της έρευνας περισσότερες από μία φορές. Από την άλλη πλευρά η *ανεπαρκής κάλυψη* είναι η αποτυχία συμπερίληψης προσδιορισμένων μονάδων του πληθυσμού στη διεξαγωγή της έρευνας. Τέλος, στα σφάλματα κάλυψης ανήκει και το φαινόμενο της εσφαλμένης ταξινόμησης (για τη ταξινόμηση έχει γίνει αναφορά στο Κεφάλαιο 1), δηλαδή τα λάθη που γίνονται κατά τη διάρκεια ταξινόμησης των συλλεχθέντων δεδομένων.

Τα *σφάλματα μέτρησης* δημιουργούνται όταν οι απαντήσεις που δίνονται διαφέρουν από τις αληθινές. Για το τύπο αυτό λαθών μπορεί να ευθύνεται ο ερωτώμενος, ο ερευνητής, το ερωτηματολόγιο που χρησιμοποιείται, η μέθοδος συλλογής δεδομένων ή το σύστημα εγγραφής-τήρησης αρχείων (εφόσον χρησιμοποιείται). Τα σφάλματα μέτρησης μπορούν να είναι είτε τυχαία είτε να εμφανίζονται με συστηματικό τρόπο. Συνεπώς, στα σφάλματα μέτρησης συμπεριλαμβάνονται λάθη που μπορεί να οφείλονται στη σύγχυση, αδιαφορία, επιπολαιότητα, μη ειλικρίνεια των ερωτώμενων, λάθη που οφείλονται στην ανεπαρκή εκπαίδευση των ερευνητών, στη τυχόν εκ των προτέρων προσδοκία τους ως προς το περιεχόμενο των απαντήσεων, ή ακόμη σκόπιμα λάθη. Τέλος μπορεί να οφείλονται στη διατύπωση των ερωτήσεων των ερωτηματολογίων, τη σειρά ή το πλαίσιο των ερωτήσεων οι οποίες παρουσιάζονται καθώς και στη μέθοδο που χρησιμοποιείται για την συλλογή των απαντήσεων.

Τα *σφάλματα μη απόκρισης* προκαλούνται όταν η έρευνα αποτυγχάνει να συλλέξει απαντήσεις για μία ή και όλες τις ερωτήσεις του ερωτηματολογίου. Το γεγονός αυτό συμβαίνει όταν αποτυγχάνεται η συλλογή όλων των απαντήσεων από όλα τα μέλη του δείγματος. Τα σφάλματα μη απόκρισης διαχωρίζονται στα σφάλματα που οφείλονται σε μονάδες που δεν αποκρίνονται (unit non-response) και στα σφάλματα που αφορούν στοιχεία για τα οποία δεν υπάρχει απόκριση (item non-response). Τα σφάλματα μη απόκρισης που αφορούν τη μη απόκριση των μονάδων εμφανίζονται όταν δεν ανταποκριθούν στην έρευνα όλα τα μέλη που ανήκουν στο δείγμα που έχει επιλεγεί. Τα σφάλματα μη απόκρισης σε στοιχεία είναι τα σφάλματα που προκαλούνται από τη μη απόκριση σε συγκεκριμένες ερωτήσεις, προκαλούν αύξηση των τυπικών σφαλμάτων (λόγω της μείωσης του αποτελεσματικού μεγέθους του δείγματος ή της χρησιμοποίησης της μεθόδου του καταλογισμού (η μέθοδος του καταλογισμού έχει αναφερθεί στο 1^ο Κεφάλαιο), ενώ μπορούν επίσης να προκαλέσουν μεροληψία στις απαντήσεις στην περίπτωση που οι ερωτώμενοι που ανταποκρίνονται διαφέρουν όσον αφορά το χαρακτηριστικό που ενδιαφέρει την έρευνα από τους ερωτώμενους που δεν ανταποκρίνονται.

Τα *σφάλματα επεξεργασίας* είναι τα λάθη στα τελικά αποτελέσματα της έρευνας και προκαλούνται από τη λαθεμένη εφαρμογή των σωστά σχεδιασμένων μεθόδων εφαρμογής. Τα σφάλματα επεξεργασίας περιλαμβάνουν όλα τα σφάλματα που αφορούν διαδικασίες μετά τη συλλογή των δεδομένων, καθώς και τη διαδικασία τύπωσης των ερωτηματολογίων. Τα περισσότερα λάθη που ανήκουν στα σφάλματα επεξεργασίας εμφανίζονται κυρίως στα δεδομένα των ανεξάρτητων μονάδων. Επίσης προκαλούνται κατά τη διάρκεια εφαρμογής μεθόδων και εκτιμήσεων. Για παράδειγμα, σε μία δειγματοληπτική έρευνα σφάλματα επεξεργασίας εμφανίζονται κατά την εγγραφή και μεταφορά των πληροφοριών, τη κωδικοποίηση αυτών, αλλά και κατά την πινακοποίηση των συλλεχθέντων πληροφοριών.

Εξίσου σημαντικό με την παρουσίαση των σφαλμάτων και των πηγών δημιουργίας τους στο τμήμα της ακρίβειας μίας ποιοτικής έκθεσης είναι η περιγραφή της μεθοδολογίας που χρησιμοποιήθηκε για την ολοκλήρωση της στατιστικής διαδικασίας. Η περιγραφή αυτή αποσκοπεί:

- στον έλεγχο του κατά πόσο η μεθοδολογία που χρησιμοποιήθηκε είναι σωστή, σύμφωνα με τις βέλτιστες πρακτικές που κυριαρχούν και τα διεθνώς αποδεκτά πρότυπα,
- στη δυνατότητα ο αναγνώστης να αντιληφθεί τις ποιοτικές αξιολογήσεις που αφορούν τις πηγές των λαθών.

Υπάρχουν πολλοί τρόποι με τους οποίους είναι δυνατό να παρουσιαστεί η μεθοδολογία που χρησιμοποιήθηκε. Η περιγραφή της μπορεί να γίνει σε μορφή παραρτήματος, ξεχωριστού κεφαλαίου, να προσχωθεί στα κατάλληλα τμήματα που αφορούν τα σφάλματα και τις πηγές αυτών ή με συνδυασμό των παραπάνω μεθόδων παρουσίασης. Τέλος, ένας ακόμη τρόπος παρουσίασης της μεθοδολογίας που χρησιμοποιήθηκε κατά τη στατιστική διαδικασία είναι η αναλυτική περιγραφή της σε ξεχωριστό έγγραφο στο οποίο θα υπάρχει αναφορά (προτείνεται η αναφορά αυτή να γίνεται με τη χρήση υπερσυνδέσμου). Παρ'όλα αυτά πρέπει να υπάρχει μία σύντομη περιγραφή της στο έγγραφο της ποιοτικής ανάλυσης, στο τμήμα της ακρίβειας.

Στο τμήμα της ακρίβειας μίας ποιοτικής έκθεσης προτείνεται επίσης να περιλαμβάνεται ένα υποκεφάλαιο που αφορά τη συνολική ακρίβεια της στατιστικής διαδικασίας. Πιο συγκεκριμένα, το υποκεφάλαιο αυτό θα ξεκινά με μία περιγραφή των βασικών πηγών τυχαίων και συστηματικών σφαλμάτων στα αποτελέσματα της έρευνας. Στη συνέχεια προτείνεται να γίνεται μία «συζήτηση» όλων των λαθών που παρουσιάστηκαν, τυχαίων και μη. Ιδιαίτερη έμφαση είναι σημαντικό να δοθεί στους βασικούς δείκτες της έρευνας. Προτείνεται να συμπεριληφθεί μία αξιολόγηση του συνολικού κινδύνου που προκαλούν τόσο τα τυχαία αλλά και τα συστηματικά λάθη.

Εξίσου σημαντική κρίνεται η αναφορά αξιολόγησης της μεροληψίας αλλά και των μεθόδων που χρησιμοποιήθηκαν για τη μείωση της. Το μέγεθός της μπορεί να παρουσιαστεί είτε ποσοτικά είτε ποιοτικά είτε και με τους δύο τρόπους. Η ποιοτική αξιολόγηση πρέπει να αναφέρεται στη συνήθη καθαρή μεροληψία που κυριαρχεί και τέλος να δίνεται μία διάταξη του μεγέθους της. Με άλλα λόγια να αναφέρεται ένας γενικός χαρακτηρισμός της, κατά πόσο δηλαδή είναι αμελητέα, μικρή, μεγάλη ή αναφέροντας τη μεγαλύτερη πιθανή της τιμή. Είναι φανερό, ότι στο τμήμα αυτό πρέπει να συμπεριληφθεί η βάση για την αξιολόγηση της μεροληψίας.

3.3.2. Αξιολόγηση

Στο σημείο αυτό πρέπει να προσθέσουμε ότι για να γίνει αναφορά των πληροφοριών για την ακρίβεια πρέπει να προηγηθεί η *αξιολόγησή* της. Παρόλο που υπάρχει μία πλούσια βιβλιογραφία για την αξιολόγηση της διασποράς δεν υπάρχουν πολλές πληροφορίες για την αξιολόγηση της ακρίβειας. Επίσης, δεν υπάρχει μέθοδος για τον υπολογισμό του συνολικού σφάλματος μίας εκτίμησης αλλά υπάρχουν μέθοδοι για την απόκτηση ένδειξης του συνολικού σφάλματος που υπάρχει.

Μία από τις μεθόδους αυτές είναι η μέθοδος σύγκρισης με άλλη πηγή. Δηλαδή, η σύγκριση της τιμής που εκτιμήθηκε με τη τιμή που έχει εκτιμηθεί από άλλη έρευνα. Για παράδειγμα, ο δείκτης της ανεργίας εκτιμάται συνήθως από έρευνες εργατικού δυναμικού αλλά και από έρευνες επιχειρήσεων. Οι διαφορές που προκύπτουν στις τιμές των δεικτών είναι στην πραγματικότητα διαφορές που

οφείλονται στο συνδυασμό σφαλμάτων και στις διαφορές μεταξύ των ορισμών. Η ανάλυση των διαφορών αυτών έχει στόχο να οδηγήσει σε πληροφορίες για το συνολικό σφάλμα.

Όταν υπάρχει σχέση μεταξύ δυο διαφορετικών παραμέτρων που εκτιμήθηκαν τότε μπορεί να υπάρξει έλεγχος που βασίζεται στις μελέτες συνοχής (consistency studies). Για παράδειγμα, είναι φανερό ότι ισχύουν οι παρακάτω σχέσεις:

- Ο αριθμός των αρσενικών που είναι παντρεμένοι είναι ίσος με τον αριθμό των θηλυκών που είναι παντρεμένες σύμφωνα με τον παραδοσιακό γάμο).
- Ο αριθμός των κατοικιών του έτους n είναι ο αριθμός των κατοικιών του έτους $(n-1)$ συν (+) τον αριθμό των νέων κατασκευών μείον (-) τον αριθμό των κατεδαφίσεων.
- Το εισόδημα ισούται με τις εξοικονομήσεις μείον (-) τα νέα δάνεια μείον (-) τις δαπάνες.

Συνεπώς, εφόσον υπάρχουν σχέσεις μεταξύ των παραμέτρων, αυτές πρέπει να ικανοποιούνται πλήρως από τα συλλεχθέντα δεδομένα. Οι σημαντικές αποκλίσεις απαιτούν περαιτέρω έλεγχο και αναζήτηση των πιθανών σφαλμάτων. Οι μελέτες συνοχής πρέπει να γίνονται πριν την ανακοίνωση των αποτελεσμάτων. Στην περίπτωση τώρα που δύο παράμετροι είναι ανεξάρτητα εκτιμημένοι η τυχόν ασυνέπεια μεταξύ δύο εκτιμήσεων αποτελεί την έναρξη της ανάλυσης των σφαλμάτων σε κάθε μία εκτίμηση.

Συγκεντρωτικά, για το τμήμα της ακρίβειας μιας ποιοτικής έκθεσης περιλαμβάνονται:

3. Ακρίβεια

- ✓ Αναφορά των βασικών σφαλμάτων και των πηγών τους.
- ✓ Παρουσίαση της μεθοδολογίας που χρησιμοποιήθηκε στη στατιστική διαδικασία.
- ✓ Αξιολόγηση των πηγών σφαλμάτων με ιδιαίτερη έμφαση στις εκτιμήσεις των βασικών δεικτών.
- ✓ Αξιολόγηση της μεροληψίας και των μεθόδων που χρησιμοποιήθηκαν για την μείωσή της.

3.3.3. Δειγματοληπτικές έρευνες

Τα είδη μίας δειγματοληπτικής έρευνας είναι δύο. Ο διαχωρισμός τους βασίζεται στην επιλογή του δείγματος. Πιο συγκεκριμένα, υπάρχουν οι έρευνες που βασίζονται σε τυχαία επιλογή δείγματος και οι έρευνες που βασίζονται σε μη τυχαία επιλογή δείγματος. Τα *Δειγματοληπτικά Σφάλματα*, όπως προαναφέρθηκε, δημιουργούνται λόγω της επιλογής ενός δείγματος του πληθυσμού από το σύνολο του πληθυσμού για τη διεξαγωγή της έρευνας. Συνεπώς, διαφοροποιούνται σύμφωνα με την επιλογή του δείγματος με αποτέλεσμα η αναφορά τους στο τμήμα της ακρίβειας να είναι διαφορετική. Βέβαιο είναι, ότι στην περίπτωση που η στατιστική διαδικασία

στην οποία αναφέρεται η ποιοτική έκθεση βασίζεται σε δειγματοληψία στο τμήμα της ακρίβειας πρέπει να υπάρχει ένα επιπλέον τμήμα που θα αφορά τα δειγματοληπτικά σφάλματα. Μία αρχική παρουσίαση των σφαλμάτων αυτών είναι αναγκαία καθώς και ο υπολογισμός ή η εκτίμηση του συνολικού δειγματοληπτικού σφάλματος για τις εκτιμήσεις που έγιναν.

3.3.3.1. Δειγματοληπτικά Σφάλματα που βασίζονται σε τυχαία επιλογή δείγματος

Τα δειγματοληπτικά σφάλματα πρέπει να αναφέρονται για όλες τις εκτιμημένες τιμές που προέκυψαν από μία στατιστική διαδικασία για την οποία έγινε δειγματοληψία. Όταν είναι σημαντική η εμφάνισή τους τότε πρέπει να γίνεται μία ποσοτική παρουσίαση αυτών συνοδευόμενη από την εκτιμημένη τιμή και τη διασπορά της. Υπάρχουν πολλοί τρόποι παρουσίασης που μπορούν να χρησιμοποιηθούν.

- ❖ Τυπικό σφάλμα. Το τυπικό σφάλμα δε χρησιμοποιείται συχνά διότι η ερμηνεία του δεν είναι τόσο προφανής για τους χρήστες.
- ❖ Συντελεστής Μεταβλητότητας (ισούται με το λόγο του τυπικού σφάλματος προς την αναμενόμενη τιμή). Είναι ουσιαστικά το τυπικό σφάλμα εκφρασμένο σε ποσοστό. Ο συντελεστής μεταβλητότητας αποτελεί την καταλληλότερη στατιστική για ποσοτικές μεταβλητές με μεγάλες θετικές τιμές. Δεν ενδείκνυται για εκτιμήσεις εκφρασμένες σε ποσοστά και μεταβλητές που μπορεί να λάβουν αρνητικές τιμές (π.χ. κέρδος).
- ❖ Διάστημα Εμπιστοσύνης. Η πραγματική ερμηνεία ενός *διαστήματος εμπιστοσύνης* είναι ότι σε μία σειρά κατασκευών διαστημάτων εμπιστοσύνης μιας παραμέτρου, έστω θ , με ανεξάρτητα δείγματα του αυτού μεγέθους ένα ποσοστό (συντελεστής εμπιστοσύνης $\times 100\%$) των διαστημάτων αυτών αναμένεται να περιέχουν την αληθή τιμή της θ . Στις περισσότερες περιπτώσεις που είναι λογικό να υποθεθεί ότι η εκτιμημένη τιμή ακολουθεί την κανονική κατανομή το διάστημα που είναι $\pm 2 \times$ εκτιμημένο τυπικό σφάλμα οδηγεί σε ένα περίπου 95% διάστημα εμπιστοσύνης. Το διάστημα μπορεί να εκφραστεί και σε ποσοστό χρησιμοποιώντας αντί του τυπικού σφάλματος το συντελεστή διακύμανσης. Τα διαστήματα εμπιστοσύνης χρησιμοποιούνται για τους βασικούς δείκτες αφού αποτελούν το πιο αυστηρό και καθαρό τρόπο παρουσίασης της μεταβλητότητας λόγω δειγματοληψίας. Χρησιμοποιούνται επίσης και στις περιπτώσεις όπου ο συντελεστής μεταβλητότητας δεν είναι κατάλληλο μέσο παρουσίασης.

Ένα ακόμη σημαντικό τμήμα που αφορά τα δειγματοληπτικά σφάλματα είναι η αναφορά των ακραίων τιμών. Ακραίες τιμές αποτελούν οι τιμές που έχουν μικρή πιθανότητα εμφάνισης. Σε μία ποιοτική έκθεση είναι σημαντικό να αναφέρεται κατά πόσο, για ποιο λόγο και σε τι βαθμό χρησιμοποιήθηκε ειδική επεξεργασία των ακραίων τιμών κατά την επεξεργασία δεδομένων.

Η μεταβλητότητα μιας τιμής που πρόκειται να υπολογιστεί μπορεί να εκφραστεί είτε με τη διασπορά, το τυπικό σφάλμα, το συντελεστή μεταβλητότητας είτε με ένα διάστημα εμπιστοσύνης για τη τιμή αυτή. Από την άλλη πλευρά το

συνολικό σφάλμα (περιέχει πληροφορίες τόσο για τη διασπορά όσο και για τη μεροληψία μιας εκτιμημένης τιμής) εκφράζεται ως η ρίζα του μέσου τετραγωνικό σφάλματος (RMSE) όπου για μία εκτιμημένη τιμή $\hat{\theta}$ το μέσο τετραγωνικό σφάλμα υπολογίζεται ως εξής :

$$MSE(\hat{\theta}) = E[(\hat{\theta} - \theta)^2]$$

και ισούται με: $MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias(\hat{\theta}, \theta)^2$, όπου $Var(\hat{\theta})$ είναι η διασπορά που υπολογίστηκε για την εκτιμημένη παράμετρο θ , και $Bias(\hat{\theta}, \theta)$ είναι η μεροληψία της εκτιμήτριας.

Παρόλο που το μέσο τετραγωνικό σφάλμα φαίνεται ιδανικό για τον άμεσο υπολογισμό της ακρίβειας μίας εκτιμημένης τιμής στην πραγματικότητα σπάνια μπορεί να υπολογιστεί. Επιπρόσθετα, ο υπολογισμός της ακρίβειας προτείνεται να γίνεται με μία μη άμεση προσέγγιση η οποία θα βασίζεται σε όλα τα μη δειγματοληπτικά σφάλματα που έγιναν ανάλογα με τη έρευνα που έλαβε χώρα.

3.3.3.2. Δειγματοληπτικά Σφάλματα που δε βασίζονται σε τυχαίο δείγμα

Στην περίπτωση των δειγματοληπτικών ερευνών οι οποίες δε βασίζονται σε τυχαίο δείγμα ο υπολογισμός του τυπικού σφάλματος είναι αρκετά πολύπλοκος και εξαρτάται από το είδος της δειγματοληψίας που χρησιμοποιήθηκε. Ο υπολογισμός βασίζεται στο μοντέλο δειγματοληψίας για το οποίο πρέπει να γίνει αναφορά καθώς επίσης και αναφορά των λόγων που οδήγησαν στη χρήση του συγκεκριμένου μοντέλου. Επίσης, σημαντική θεωρείται μία «συζήτηση» για τη δειγματοληπτική μεροληψία, δηλαδή την εμφάνιση μεροληψίας λόγω της μη τυχαίας δειγματοληψίας. Περισσότερες πληροφορίες δίνονται στα εγχειρίδια μιας ποιοτικής έκθεσης. Συνολικά έχουμε ότι στην περίπτωση δειγματοληπτικών ερευνών το τμήμα της ακρίβειας πρέπει να περιέχει τις παρακάτω επιπρόσθετες πληροφορίες:

Δειγματοληπτικά Σφάλματα

- ✓ Παρουσίαση των δειγματοληπτικών σφαλμάτων της στατιστικής διαδικασίας.
- ✓ Εφόσον είναι εφικτό, αναφορά του συνολικού δειγματοληπτικού σφάλματος για τις εκτιμήσεις μεταβλητών.

Τυχαία Δειγματοληψία

- ✓ Αναφορά των υπολογισμένων δειγματοληπτικών σφαλμάτων και των μεθόδων υπολογισμού τους.
- ✓ Αναφορά των ακραίων τιμών και της ειδικής επεξεργασίας τους.

Μη Τυχαία Δειγματοληψία

- ✓ Εκτίμηση του τυπικού σφάλματος μέσω ενός μοντέλου δειγματοληψίας. Αναφορά του μοντέλου και του λόγου επιλογής του.
- ✓ «Συζήτηση» για την εμφάνιση δειγματοληπτικής μεροληψίας.

3.4. Επικαιρότητα και Εγκαιρότητα (Timeliness and Punctuality)

Η *Επικαιρότητα* των στατιστικών αποτελεσμάτων είναι το διάστημα του χρόνου που υπάρχει μεταξύ του φαινομένου ή του συμβάντος στο οποίο αναφέρονται τα αποτελέσματα και του χρόνου διαθεσιμότητας τους. Από την άλλη πλευρά η έννοια της *Εγκαιρότητας* αναφέρεται στο χρόνο καθυστέρησης ανάρτησης των αποτελεσμάτων σε σχέση με την ημερομηνία που ζητήθηκαν να παραδοθούν. Η ημερομηνία παράδοσης καταγράφεται σε επίσημο ημερολόγιο και έχει προσυμφωνηθεί από τους συνεργάτες.

Η Επικαιρότητα είναι σχετικά εύκολη και άμεση στον υπολογισμό της. Ένας τρόπος υπολογισμού της είναι η μέτρηση του χρόνου μεταξύ της τελευταίας ημερομηνίας στην οποία αναφέρονται και στην ημερομηνία διάθεσής τους.

Η παρουσίαση της εγκαιρότητας είναι παρομοίως απλή. Υπολογίζεται ως το ποσοστό των παραδοθέντων αποτελεσμάτων στην ώρα τους σύμφωνα με τον κανονισμό που ισχύει. Ο κανονισμός αυτός υπάρχει είτε σε επίσημα χρονοδιαγράμματα ή σε άλλες συμφωνίες. Η περίπτωση αυτή αφορά τις χρόνιες ή και πιο συχνές διαθέσεις αποτελεσμάτων.

Κάποιες στατιστικές διαδικασίες παραδίδονται σε διάφορες εκδόσεις, για παράδειγμα την προκαταρκτική έκδοση, αναθεωρημένη ή τελική. Στην περίπτωση αυτή κάθε παράδοση έχει τη δική της επικαιρότητα και εγκαιρότητα. Οι διαφορετικές εκδόσεις πρέπει να διαφοροποιηθούν και να αναφερθούν ξεχωριστά στο τμήμα αυτό της ποιοτικής έκθεσης.

Τέλος, είναι σημαντικό να αναφέρονται οι λόγοι τόσο της μεγάλης τιμής του διαστήματος χρόνου που περιγράφεται στην επικαιρότητα όσο και του διαστήματος που περιγράφεται από την έννοια εγκαιρότητα.

Συνεπώς, συνολικά έχουμε ότι στο τμήμα της Επικαιρότητας και Εγκαιρότητας της ποιοτικής έκθεσης δίνονται πληροφορίες που αφορούν:

4. Επικαιρότητα και Εγκαιρότητα

- ✓ Αναφορά του διαστήματος χρόνου μεταξύ της παράδοσης των αποτελεσμάτων και του χρόνου στον οποίο αναφέρονται.
- ✓ Αναφορά του διαστήματος μεταξύ του χρόνου που ζητήθηκε και προσυμφωνήθηκε η παράδοση των αποτελεσμάτων και του χρόνου παράδοσής τους.
- ✓ Περιγραφή των λόγων που οδήγησαν στην αύξηση των παραπάνω διαστημάτων.

3.5. Προσβασιμότητα και Σαφήνεια (Accessibility and Clarity)

Προσβασιμότητα ορίζεται ως το μέτρο της ευκολίας πρόσβασης των χρηστών στα δεδομένα. Είναι καθορισμένο σύμφωνα με τις φυσικές συνθήκες που επικρατούν στα μέσα με τα οποία μπορεί ένας χρήστης να έχει πρόσβαση στα δεδομένα. Δηλαδή, έχει να κάνει με το που πρέπει να απευθυνθεί, το πως θα ζητήσει ένας χρήστης τα δεδομένα, την ημερομηνία παράδοσής τους, τις συνθήκες που αφορούν το marketing των αποτελεσμάτων (δικαίωμα αντιγραφής κ.λπ.), τη διαθεσιμότητά τους, την ποικιλία των μορφών παράδοσής τους (επίσημα έγγραφα, αρχεία, διαδίκτυο κ.λπ.).

Η *Σαφήνεια* των αποτελεσμάτων ορίζεται ως το μέτρο ευκολίας αντίληψης των δεδομένων από τους χρήστες. Καθορίζεται σύμφωνα με τις πληροφορίες που συνοδεύουν τα δεδομένα που παρουσιάζονται (metadata). Πιο συγκεκριμένα η έννοια σαφήνεια αφορά το βαθμό στον οποίο τα δεδομένα συνοδεύονται από κατάλληλες πληροφορίες, το βαθμό στον οποίο υπάρχει απεικόνιση των δεδομένων (γραφήματα, απεικονίσεις κ.λπ.), τη διαθεσιμότητα πληροφοριών που αφορούν την ακρίβεια των αποτελεσμάτων, το εύρος της επιπλέον βοήθειας που διατίθεται από την παραγωγή των δεδομένων.

Η αξιολόγηση της προσβασιμότητας παίρνει διάφορες μορφές αφού εξαρτάται από α) τη διάδοση των καναλιών στα οποία αναρτώνται τα αποτελέσματα, β) τη φόρμα στην οποία παραδίδονται (μικροδεδομένα-microdata, συνολικά δεδομένα) γ) τη πολιτική τιμολόγησής τους. Στο τμήμα της ποιοτικής έκθεσης που αφορά την προσβασιμότητα είναι σημαντικό να αναφέρονται τα διάφορα μέσα παράδοσης των αποτελεσμάτων (εφημερίδα, διαδίκτυο κ.λπ.), η πολιτική τιμολόγησής τους, οι επιπτώσεις στην πρόσβαση των αποτελεσμάτων αλλά και τα όρια που τίθενται λόγω της εμπιστευτικότητας των δεδομένων.

Από την άλλη πλευρά, η σαφήνεια των αποτελεσμάτων εξαρτάται από τον όγκο και την ποιότητα των πληροφοριών που συνοδεύουν τα δεδομένα. Γι αυτό το λόγο στην ποιοτική έκθεση κρίνεται σημαντικό να αναφέρεται μία σύντομη περιγραφή των διαθέσιμων αυτών πληροφοριών.

Σε επίσημο έγγραφο του 2008 που αφορά τα μέτρα της προσβασιμότητας και της σαφήνειας τονίζονται σημαντικά σημεία που αφορούν τα τμήματα αυτά. Η

περιγραφή βασίζεται στη διαίρεση των χρηστών σε κατηγορίες(π.χ. περιστασιακούς χρήστες, επαγγελματίες χρηστές κ.λ.π.)¹⁰.

Επίσης, η καλύτερη πηγή εκτίμησης της σαφήνειας και του επιπέδου της προσβασιμότητας των δημοσιευμένων δεδομένων θεωρήθηκε η ανατροφοδότηση από τους χρήστες. Με την έννοια αυτή, ερωτήσεις προς τους χρήστες που αφορούν την εμπειρία τους, πάνω στα δύο θέματα της σαφήνειας και της προσβασιμότητας, είναι σημαντικό να περιέχονται συνοδευόμενες και από τις απαντήσεις τους. Τόσο η αναφερθείσα όσο και οποιαδήποτε άλλη ανατροφοδότηση από τους χρήστες είναι σημαντικό να αναφέρονται στο τμήμα αυτό.

Τελευταίο και σημαντικό μέρος του τμήματος αυτού της ποιοτικής ανάλυσης κρίνεται αναγκαίο να είναι η αναφορά των πρόσφατων καθώς και των σχεδιαζόμενων βελτιώσεων που αφορούν το επίπεδο της προσβασιμότητας και της σαφήνειας.

Συνεπώς συνολικά για το τμήμα της προσβασιμότητας και της σαφήνειας έχουμε:

5. Προσβασιμότητα και Σαφήνεια

- ✓ Περιγραφή των επικρατούσων συνθηκών που αφορούν την προσβασιμότητα στα δεδομένα (ΜΜΕ, υποστήριξη, πολιτική τιμολόγησης, πιθανούς περιορισμούς κ.α.).
- ✓ Περίληψη των πληροφοριών που συνοδεύουν τα δεδομένα (metadata) (επίσημα έγγραφα, επεξηγήσεις, ποιοτικά όρια κ.α.).
- ✓ Η περιγραφή πρέπει να αφορά όλους τους χρήστες και το πώς οι ανάγκες τους συνυπολογίστηκαν.
- ✓ Περίληψη της ανατροφοδότησης από τους χρήστες για την προσβασιμότητα και τη σαφήνεια των δεδομένων.

3.6. Συνοχή και Συγκρισιμότητα (Coherence and Comparability)

Η *συνοχή* δύο ή περισσότερων αποτελεσμάτων ορίζεται ως ο βαθμός στον οποίο οι στατιστικές διαδικασίες από τις οποίες προήλθαν τα αποτελέσματα χρησιμοποιούσαν κοινό σκεπτικό και εναρμονισμένες μεταξύ τους μεθόδους. Με την έννοια σκεπτικό αναφερόμαστε στην ταξινόμηση, τους ορισμούς και τον πληθυσμό προς μελέτη. Το πλεονέκτημα των *συνεκτικών* αποτελεσμάτων είναι ότι τα αποτελέσματα των δύο διαδικασιών μπορούν να συνδυαστούν και να χρησιμοποιηθούν από κοινού. Δύο περιπτώσεις από κοινού χρησιμοποίησης τέτοιων αποτελεσμάτων υπάρχουν. Η περίπτωση που τα αποτελέσματα αναφέρονται στον ίδιο πληθυσμό, χρονική περίοδο και περιοχή αλλά περιλαμβάνουν διαφορετικά δεδομένα (π.χ. δεδομένα απασχόλησης και στοιχεία παραγωγής). Επίσης, από κοινού χρησιμοποίηση των αποτελεσμάτων έχουμε και στην περίπτωση που τα

¹⁰ <http://unstats.un.org/unsd/accsub/2008docs-CDQIO/Ses3-Pap3.pdf>

αποτελέσματα αναφέρονται σε διαφορετικές περιοχές ή χρονικές περιόδους αλλά περιλαμβάνουν τα ίδια δεδομένα (π.χ. δεδομένα απασχόλησης).

Η *συγκρισιμότητα* των αποτελεσμάτων είναι μία ειδική περίπτωση της συνοχής των αποτελεσμάτων και πιο συγκεκριμένα η περίπτωση που αναφέρθηκε όπου τα αποτελέσματα αναφέρονται στα ίδια στατιστικά δεδομένα αλλά στοχεύουν σε χρονική σύγκριση, σύγκριση μεταξύ περιοχών ή άλλων συγκρίσεων.

Γνωρίζουμε ότι, διαφορετικές ομάδες δεδομένων (π.χ., δεδομένα απασχόλησης) συλλέγονται από διαφορετικές στατιστικές διαδικασίες. Η έννοια της συνοχής χρησιμοποιείται όταν είναι εφικτή η αξιολόγηση του κατά πόσο τα διαφορετικά στατιστικά αποτελέσματα είναι αξιόπιστα όταν συνδυαστούν. Από την άλλη πλευρά, η έννοια της συγκρισιμότητας χρησιμοποιείται όταν μπορεί να αξιολογηθεί το κατά πόσο αποτελέσματα από την ίδια στατιστική διαδικασία αλλά αναφερόμενα σε διαφορετικές χρονικές περιόδους ή περιοχές ή άλλη πηγή είναι αξιόπιστο να χρησιμοποιηθούν από κοινού. Πιο συγκεκριμένα, από τις δύο περιπτώσεις που αναφέρθηκαν για την από κοινού χρησιμοποίηση δεδομένων, η ισχύς της από κοινού χρήσης των δεδομένων απασχόλησης και των στοιχείων παραγωγής που αναφέρονται στο ίδιο πληθυσμό, χρονική περίοδο και περιοχή εξαρτάται από τη συνοχή των αποτελεσμάτων ενώ αντίθετα στην περίπτωση χρήσης ίδιων δεδομένων που αναφέρονται στον ίδιο πληθυσμό και περιοχή αλλά σε διαφορετική χρονική περίοδο εξαρτάται από τη συγκρισιμότητα των αποτελεσμάτων.

Είναι αξιοσημείωτο το γεγονός ότι παρόλο που οι έννοιες συνοχή και συγκρισιμότητα αφορούν τα στατιστικά αποτελέσματα εξαρτώνται και αξιολογούνται από τις στατιστικές διαδικασίες που χρησιμοποιούνται για την παραγωγή των αποτελεσμάτων. Επίσης, οι έννοιες της ακρίβειας και της συνοχής/συγκρισιμότητας συχνά συγχέονται. Η συνοχή/συγκρισιμότητα αναφέρεται και υπολογίζεται σύμφωνα με τις πληροφορίες που συνοδεύουν τα δεδομένα και αφορούν το σκεπτικό και τις μεθόδους των διαδικασιών ενώ η ακρίβεια πηγάζει από τις πληροφορίες που συνοδεύουν τα δεδομένα και αφορούν τη συχνότητα δειγματοληψίας (sampling rates), τα ποσοστά σφάλματος στη συλλογή δεδομένων κ.α..

Στο σημείο αυτό είναι σημαντικό να δώσουμε κάποιες περιπτώσεις συνοχής/συγκρισιμότητας καθώς και πιθανούς λόγους απώλειας της συνοχής ή της συγκρισιμότητας.

3.6.1. Περιπτώσεις συνοχής/συγκρισιμότητας

- Συγκρισιμότητα με την πάροδο του χρόνου, για παράδειγμα σύγκριση μεταξύ των αποτελεσμάτων διαφορετικών μηνών, ετών.
- Συγκρισιμότητα ανά περιοχή, για παράδειγμα σύγκριση των αποτελεσμάτων που αφορούν τον ίδιο μήνα για διαφορετικά κράτη – μέλη της Ευρωπαϊκής Ένωσης.
- Άλλες περιπτώσεις συγκρισιμότητας. Αφορούν τις περιπτώσεις όπου επιτυγχάνεται σύγκριση σε θέματα που προκύπτουν π.χ. ανά οικονομική

δραστηριότητα μίας ομάδας, ανά επαγγελματική ομάδα ή μεταξύ των δύο φύλων.

- Εσωτερική Συνοχή. Αφορά τα αποτελέσματα που προκύπτουν από μία μοναδική στατιστική διαδικασία η οποία όμως αποτελείται από διαφορετικά τμήματα για μία χρονική περίοδο και περιοχή.
- Συνοχή μεταξύ ετήσιων και μη ετήσιων (μικρότερων χρονικών περιόδων από ένα έτος) στατιστικών διαδικασιών, για παράδειγμα ετήσια και μηνιαία αποτελέσματα για τα στοιχεία παραγωγής για ίδιες βιομηχανίες στην ίδια περιοχή.
- Συνοχή με τους Εθνικούς Λογαριασμούς. Αφορά τις οικονομικές έρευνες που εφοδιάζουν τους Εθνικούς Λογαριασμούς.
- Συνοχή με άλλες στατιστικές διαδικασίες. Για παράδειγμα συνοχή μεταξύ της απασχόλησης που εκτιμάται από μία έρευνα και του αριθμού των εργαζομένων που εκτιμήθηκε από μία άλλη έρευνα.

3.6.2. Λόγοι Έλλειψης Συνοχής/Συγκρισιμότητας

Οι βασικοί λόγοι οι οποίοι μπορούν να οδηγήσουν στην έλλειψη συνοχής/συγκρισιμότητας είναι οι διαφορές στο σκεπτικό και τις μεθόδους που χρησιμοποιούνται. Οι διαφορές προκύπτουν από τις αλλαγές με την πάροδο του χρόνου στη στατιστική διαδικασία που εφαρμόζεται. Οι αλλαγές αυτές μπορεί να προκύψουν από διάφορες αιτίες, π.χ. βελτίωση των ερωτηματολογίων, μέθοδοι, αυτοματισμός, νέα τεχνολογία, πιο σύγχρονες ταξινομήσεις, αλλαγή νομοθεσίας, χρηματοδότησης με αποτέλεσμα την αλλαγή του δείγματος κ.α. Οι πιθανοί λόγοι έλλειψης συνοχής/συγκρισιμότητας κατηγοριοποιούνται ως εξής:

Στους λόγους που αφορούν διαφορές στο σκεπτικό:

- Πληθυσμός-Μέλη και Κάλυψη. Οι πληθυσμοί που στοχεύετε να μελετηθούν είναι πιθανό να διαφέρουν μεταξύ τους.
- Γεωγραφική Κάλυψη. Αγροτικές περιοχές μπορεί να περιλαμβάνονται στη μία έρευνα ενώ σε μία άλλη όχι.
- Χρονική Περίοδος. Για παράδειγμα μία ετήσια έρευνα μπορεί να αναφέρεται στο έτος που ξεκινά το Μάρτιο ενώ μία άλλη στο έτος που ξεκινά το Σεπτέμβριο (ημερολογιακό έτος).
- Ορισμοί επεξηγήσεις των δεδομένων, ταξινόμηση.

Αφορούν τις μεθόδους που χρησιμοποιούνται:

- Το πλαίσιο του πληθυσμού.
- Πηγές των δεδομένων και σχεδιασμός του δείγματος.
- Συλλογή των δεδομένων, συγκέντρωση και διόρθωση.
- Καταλογισμός πεδίων και εκτίμηση.

Αυτές είναι οι κύριες κατηγορίες από τις οποίες πηγάζει η έλλειψη συνοχής/συγκρισιμότητας. Οι πληροφορίες που δόθηκαν κρίθηκαν απαραίτητες για την σωστή κατανόηση των δύο εννοιών.

Στο τμήμα της συνοχής/συγκρισιμότητας μίας ποιοτικής έκθεσης είναι σημαντικό να περιλαμβάνεται μία γενική προσέγγιση του θέματος. Με την έννοια αυτή, να αναφέρονται οι λόγοι που προκάλεσαν την έλλειψη συνοχής/συγκρισιμότητας και να επεξηγούνται πλήρως. Επιπλέον, κρίνεται απαραίτητη η ποσοτική εκτίμηση των επιπτώσεων των βασικών πηγών έλλειψης συνοχής/συγκρισιμότητας. Στους λόγους που προκάλεσαν την τυχόν έλλειψη σημαντικό θα ήταν να περιλαμβάνονται και οι αλλαγές στη στατιστική διαδικασία που συντέλεσαν στην έλλειψη αυτή. Ως εκ τούτου προτείνεται να δίνεται ένας πίνακας στον οποίο να αναφέρονται ξεχωριστά όλοι οι λόγοι που συντέλεσαν στην έλλειψη, η ποσοτική εκτίμηση των επιπτώσεων της κάθε περίπτωσης και στη συνέχεια η συνολική επίπτωση όλων των περιπτώσεων στη συνοχή/συγκρισιμότητα.

Όσον αφορά τη συγκρισιμότητα με την πάροδο του χρόνου αποτελεί ζωτικής σημασίας ποιοτικό ζήτημα όλων των στατιστικών αποτελεσμάτων που δημοσιεύονται διαδοχικά. Για τους περισσότερους χρήστες, ο εντοπισμός των αλλαγών σε κοινωνικά και οικονομικά ζητήματα με την πάροδο του χρόνου είναι υψίστης σημασίας με αποτέλεσμα η συγκρισιμότητα με την πάροδο του χρόνου να είναι απαραίτητη. Συνεπώς, οι χρήστες πρέπει να ενημερώνονται για τους περιορισμούς που υπάρχουν ως προς τη χρήση των δεδομένων για σύγκριση με την πάροδο του χρόνου. Αυτό γίνεται μέσα από την ποιοτική έκθεση. Με την έννοια αυτή, πρέπει να καθορίζεται η έκταση των αλλαγών στη στατιστική διαδικασία από τη μία περίοδο στην επόμενη, δηλαδή, κατά πόσο έχουν γίνει αλλαγές και στην περίπτωση που έχουν γίνει η καταγραφή του βαθμού στον οποίο επηρεάζουν το σχεδιασμό της χρονοσειράς.

Επίσης, στο τμήμα της συνοχής/συγκρισιμότητας της ποιοτικής έκθεσης είναι σημαντικό να αναφέρεται η εκτίμηση της συγκρισιμότητας ανά περιοχή. Η αναφορά αυτή μπορεί να γίνει με δύο τρόπους, είτε με την σύγκριση ανά ζευγάρια των πληροφοριών που δίνονται και συνοδεύουν τα δεδομένα ανά περιοχές, είτε με τη σύγκριση των πληροφοριών που συνοδεύουν τα δεδομένα με μία περιοχή που ακολουθεί τα πρότυπα του Εθνικού Στατιστικού Συστήματος. Υπάρχουν δύο καταστάσεις από τις οποίες η ισχύουσα είναι σημαντικό να αναφέρεται. Πιο συγκεκριμένα, είτε χρησιμοποιείται η ίδια στατιστική διαδικασία ανάμεσα στις δύο περιοχές είτε χρησιμοποιείται διαφορετική διαδικασία.

Ομοίως για οποιαδήποτε άλλη δυνατότητα συγκρισιμότητας των αποτελεσμάτων (αναφέρθηκαν οικονομική δραστηριότητα μίας ομάδας χρηστών, επαγγελματική ομάδα ή μεταξύ των δύο φύλων) είναι σημαντικό να αναφέρεται αν χρησιμοποιείται το ίδιο στατιστικό εργαλείο ή όχι.

Όσον αφορά τη συνοχή των ετήσιων και μη ετήσιων αποτελεσμάτων οι λόγοι έλλειψης συνοχής (εφόσον υπάρχουν) κρίνεται αναγκαίο να αναφέρονται και να εξηγούνται πλήρως. Μία άμεση μέθοδος για την αξιολόγηση του μεγέθους των διαφορών που προκαλούνται λόγω αυτής της έλλειψης είναι η σύγκριση των ετήσιων και των μικρότερων από ετήσιων αποτελεσμάτων (π.χ. αν τόσο τα ετήσια όσο και τα μηνιαία αποτελέσματα μετρούν επίπεδα τότε θα πρέπει από τα μηνιαία αποτελέσματα να προκύπτουν τα χρόνια και να συγκρίνονται με τα ετήσια αποτελέσματα).

Η έλλειψη συνοχής των δεδομένων που παρελήφθησαν από διαφορετικές στατιστικές διαδικασίες μπορεί εύκολα να εντοπιστεί με μία διαδικασία κατάρτισης από τους Εθνικούς Λογαριασμούς ανεξαρτήτως αν πρόκειται για άμεσες έρευνες ή έρευνες βασιζόμενες σε μητρώα. Συνεπώς, η έλλειψη συνοχής με τους Εθνικούς Λογαριασμούς μπορεί εύκολα να μετρηθεί.

Τέλος, μπορεί να υπάρχουν άλλες στατιστικές οι οποίες μπορούν να χρησιμοποιηθούν σε συνδυασμό για την παραγωγή αποτελεσμάτων. Όπως έχει ήδη αναφερθεί στην ποιοτική έκθεση είναι απαραίτητο να αναφέρεται η αξιολόγηση της έλλειψης συνοχής μεταξύ των αποτελεσμάτων που συνδυάζονται καθώς και οι επιπτώσεις της.

Συνολικά, το τμήμα της Συνοχής και Συγκρισιμότητας μίας ποιοτικής έκθεσης είναι σημαντικό να περιλαμβάνει τα εξής:

6. Συνοχή και Συγκρισιμότητα

- ✓ Γενική προσέγγιση. Περιγραφή όλων των σκεπτικών και μεθόδων που επηρέασαν τη συνοχή/συγκρισιμότητα, αξιολόγηση των πιθανών επιπτώσεων κάθε αναφερόμενης διαφοράς στα αποτελέσματα.
- ✓ Αναφορά περιόδων όπου δεν υπάρχει συνέχεια της χρονοσειράς των αποτελεσμάτων, ο λόγος και οι μέθοδοι αντιμετώπισης.
- ✓ Ποσοτική αξιολόγηση της συγκρισιμότητας ανά περιοχές βασισμένη στον αριθμό των διαφορών που αναφέρονται στις πληροφορίες που συνοδεύουν τα δεδομένα.
- ✓ Αναφορά των πιθανών πηγών έλλειψης συγκρισιμότητας που υπολογίζονται σύμφωνα με διεθνή πρότυπα.
- ✓ Αναφορά οποιασδήποτε έλλειψης συνοχής στα αποτελέσματα της στατιστικής διαδικασίας.
- ✓ Το αποτέλεσμα της σύγκρισης των αποτελεσμάτων με τους Εθνικούς Λογαριασμούς.
- ✓ Εφόσον τα στατιστικά αποτελέσματα συνδυάζονται με τα αποτελέσματα άλλων διαδικασιών, αναφορά των περιορισμών που θέτονται για τη διατήρηση της συνοχής.

3.7. Συμβιβασμοί μεταξύ των ποιοτικών συνιστωσών των αποτελεσμάτων.

Όπως έχει γίνει φανερό και έχει αναφερθεί δεν υπάρχει σαφής διαχωρισμός μεταξύ των συνιστωσών των τμημάτων μίας ποιοτικής έκθεσης. Στο τμήμα αυτό αναφέρονται οι συμβιβασμοί που έχουν γίνει και αφορούν τις περιπτώσεις όπου η βελτίωση μία συνιστώσας ενός τμήματος σχετίζεται με την επιδείνωση μίας συνιστώσας ενός άλλου τμήματος της ποιοτικής έκθεσης. Αναφέρονται οι σημαντικότερες περιπτώσεις όπου χρειάζεται συμβιβασμός.

1. Μεταξύ της Συνάφειας και της Ακρίβειας.
2. Μεταξύ της Συνάφειας και της Επικαιρότητας.
3. Μεταξύ της Συνάφειας και της Συνοχής.

4. Μεταξύ της Συνάφειας και της Συγκρισιμότητας κατά την πάροδο του χρόνου.
5. Μεταξύ της Συγκρισιμότητας ανά περιοχή και της Συγκρισιμότητας κατά την πάροδο του χρόνου.
6. Μεταξύ της Ακρίβειας και της Επικαιρότητας.

Συνοπτικά στο τμήμα αυτό της ποιοτικής έκθεσης αναφέρονται:

7. Συμβιβασμοί μεταξύ των ποιοτικών συνιστωσών των αποτελεσμάτων.
- ✓ Αναφορά των πιο σημαντικών συμβιβασμών που έγιναν και το πλαίσιο στο οποίο βασίστηκαν.
 - ✓ Αναφορά περίπτωσης όπου έπρεπε να τεθεί συμβιβασμός και ο λόγος για τον οποίο αυτό δεν έγινε.

3.8. Εκτίμηση των αναγκών και των αντιλήψεων των χρηστών.

Η εκτίμηση των αναγκών και των αντιλήψεων των χρηστών περιέχει φυσικά την εξέταση των αποτελεσμάτων των ποιοτικών συνιστωσών της ποιοτικής έκθεσης. Το τμήμα αυτό της ποιοτικής έκθεσης περιέχει όλες τις πτυχές εκτίμησης των αναγκών και αντιλήψεων των χρηστών που δεν περιορίζονται στην απλή εξέταση των αποτελεσμάτων των συνιστωσών της ποιοτικής έκθεσης.

Η αξιολόγηση των αναγκών των χρηστών δεν είναι εύκολη υπόθεση αφού υπάρχουν πολλοί χρήστες των αποτελεσμάτων μιας στατιστικής διαδικασίας καθώς και πολλές διαφορετικές χρήσεις των αποτελεσμάτων από τους χρήστες. Συνεπώς το πρώτο βήμα είναι η συλλογή πληροφοριών για τους χρήστες-ποιοί είναι, τον αριθμό τους και πόσο σημαντικοί είναι μεμονωμένα και συνολικά. Κατασκευάζεται, με αυτό τον τρόπο, μία ταξινόμηση των χρηστών σε ομάδες και ο υπολογισμός του μεγέθους της κάθε ομάδας. Δεύτερο βήμα αποτελεί ο καθορισμός των αναγκών των χρηστών της κάθε ομάδας και στην περίπτωση σημαντικών αναγκών μεμονωμένων χρηστών η αναφορά τους. Το τρίτο βήμα είναι ο ορισμός προτεραιοτήτων όσον αφορά τις διαφορετικές ανάγκες κάθε ομάδας χρηστών. Ένα προφανές παράδειγμα είναι ότι οι ανάγκες των χρηστών που αναπτύσσουν την πολιτική μίας χώρας προηγούνται των αναγκών των ακαδημαϊκών ερευνητών. Επίσης, ένα τέταρτο βήμα είναι ο καθορισμός των πληροφοριών που συνοδεύουν τα δεδομένα, δηλαδή τι υλικό ποιοτικής επεξήγησης πρέπει να παραδοθεί και σε ποια μορφή.

Συνολικά, φαίνεται ότι στο τμήμα αυτό της ποιοτικής έκθεσης, όσον αφορά την εκτίμηση των αναγκών των χρηστών, συμπεριλαμβάνεται η ταξινόμηση των χρηστών, ο εντοπισμός των χρήσεων των αποτελεσμάτων της στατιστικής διαδικασίας, ο καθορισμός προτεραιοτήτων ανάμεσα στις ανάγκες των χρηστών, η αναφορά των διαδικασιών που συλλέχθηκαν οι πληροφορίες αυτές.

Από την άλλη πλευρά όσον αφορά την εκτίμηση των αντιλήψεων των χρηστών, γνωρίζοντας ότι η ικανοποίηση των χρηστών είναι υψίστης σημασίας για μία στατιστική διαδικασία, η πιο αποτελεσματική μέθοδος υπολογισμού των αντιλήψεων των χρηστών είναι η σύνταξη μίας έρευνας που αφορά την ικανοποίηση

των χρηστών. Η σύνταξη μίας νέα έρευνας δεν είναι πάντα εφικτή κυρίως λόγω οικονομικών προβλημάτων και έλλειψης χρόνου. Ωστόσο, υπάρχουν μέθοδοι εκτίμησης της ικανοποίησης των χρηστών όπως είναι η ανάλυση των πωλήσεων δημοσίευσης των αποτελεσμάτων, των σχολίων των χρηστών, των αιτήσεων και των παραπόνων που λαμβάνονται καθώς επίσης και η ανάλυση πρόσβασης στην ιστοσελίδα ανάρτησης των αποτελεσμάτων αλλά και άλλες μέθοδοι.

Συνεπώς, στο τμήμα αυτό της ποιοτικής έκθεσης είναι σημαντικό να περιλαμβάνονται επίσης τα βασικά αποτελέσματα που αφορούν την ικανοποίηση των χρηστών, τα οποία βέβαια προτείνεται να διαχωρίζονται ανάλογα με τη ταξινόμηση των χρηστών. Επίσης, σημαντικό είναι να αναφέρονται οι μέθοδοι που χρησιμοποιήθηκαν για τον υπολογισμό της ικανοποίησης των χρηστών αλλά και τυχόν μέθοδοι που εφαρμόστηκαν για τη βελτίωση της.

Άρα το τμήμα αυτό συνολικά κρίνεται αναγκαίο να περιέχει τις παρακάτω πληροφορίες:

8. Εκτίμηση των αναγκών και των αντιλήψεων των χρηστών

- ✓ Μέθοδοι και μέσα που χρησιμοποιήθηκαν για τη συλλογή πληροφοριών για τους χρήστες και τις χρήσεις των αποτελεσμάτων.
- ✓ Περιγραφή και Ταξινόμηση των χρηστών.
- ✓ Χρήσεις των αποτελεσμάτων από τους χρήστες.
- ✓ Χρήστες και χρήσεις για τις οποίες δόθηκε ιδιαίτερη σημασία.
- ✓ Μέσα με τα οποία συλλέχθηκαν πληροφορίες για την άποψη των χρηστών.
- ✓ Βασικά αποτελέσματα που αφορούν το βαθμό ικανοποίησης των χρηστών.
- ✓ Καταγραφή της πιο πρόσφατης ημερομηνίας σύνταξης έρευνας που αφορά την ικανοποίηση των χρηστών.

3.9. Επιδόσεις, Κόστος και Επιβάρυνση Ανταπόκρισης

Οι διαθέσιμοι πόροι οποιασδήποτε στατιστικής διαδικασίας είναι σημαντικό να χρησιμοποιούνται με τέτοιο τρόπο ώστε να επιτυγχάνεται η βέλτιστη χρήση τους. Αποδοτικότητα ορίζεται η μέγιστη παραγωγή αποτελεσμάτων με χρήση των πόρων που διατίθενται για τη διαδικασία. Η επιβάρυνση ανταπόκρισης η οποία ορίζεται ως η προσπάθεια-η οποία εκφράζεται σε σχέση με το χρόνο και το κόστος- που απαιτείται για την επίτευξη ικανοποιητικών απαντήσεων από τους ερωτώμενους σε μία έρευνα θα πρέπει να είναι ανάλογη με τις ανάγκες των χρηστών και όχι υπερβολική προς τους ερωτώμενους. Με την έννοια αυτή θα πρέπει να καταβάλλεται τόση προσπάθεια ώστε να καλύπτονται οι ανάγκες των χρηστών των αποτελεσμάτων αλλά στο βαθμό όπου δε θα θεωρηθεί υπερβολική από την πλευρά των ερωτώμενων. Είναι φανερό ότι η προσπάθεια αυτή, με άλλα λόγια η επιβάρυνση ανταπόκρισης, πρέπει να δύναται να μετρηθεί καθώς και οι στόχοι που τίθενται με σκοπό τη μείωση της.

Η δυνατότητα υπολογισμού του συνολικού κόστους είναι ουσιώδης για την αξιολόγηση της ποιότητας και των επιδόσεων μίας έρευνας. Η ανάλυση του κόστους σε σχέση με το όφελος κρίνεται αναγκαία έτσι ώστε να μπορούν να καθοριστούν οι κατάλληλοι συμβιβασμοί μεταξύ του κόστους και των ωφέλιμων στοιχείων που εκφράζονται από την ποιότητα των αποτελεσμάτων. Κόστος θεωρείται, φυσικά προς τους ερωτώμενους, η συμμετοχή τους, με αποτέλεσμα το κόστος αυτό να πρέπει να αντισταθμίζεται με τα οφέλη που προκύπτουν από τα δεδομένα που συλλέγονται από τη συμμετοχή τους στην έρευνα. Ο υπολογισμός του συνολικού κόστους που συνδέεται με τη στατιστική διαδικασία δεν είναι εύκολη υπόθεση αφού είναι αναγκαίος ένας μηχανισμός κατανομής των «κοινών» κόστων (όπως για παράδειγμα το κόστος από το μητρώο καταγραφής επιχειρήσεων) και των γενικών εξόδων (π.χ. ο αναγκαίος χώρος εργασίας που απαιτείται για μία έρευνα). Από την άλλη πλευρά ο υπολογισμός των άμεσων κόστων όπως είναι τα κόστη που προκύπτουν από την πρόσληψη προσωπικού, τη συλλογή δεδομένων, τα κόστη της επιλεγμένης διαδικασίας μπορεί να θεωρηθεί εύκολη υπόθεση.

Όσον αφορά την Επιβάρυνση Ανταπόκρισης των ερωτώμενων, είτε πρόκειται για ιδιώτες, μέλη νοικοκυριού ή επιχειρήσεις μπορεί να μετρηθεί βάσει των ωρών που ξοδεύονται ή με οικονομικούς όρους. Ένας συνήθης τρόπος υπολογισμού της Επιβάρυνσης Ανταπόκρισης είναι υπολογίζοντας το γινόμενο των αριθμών των συμπληρωμένων ερωτηματολογίων με το χρόνο που κατά μέσο όρο χρειάστηκε για τη συμπλήρωση του ενός ερωτηματολογίου. Ο μέσος χρόνος υπολογίζεται είτε από τη μελέτη αυτού για κάποιες ομάδες ερωτηθέντων (focus groups) είτε με τη καταγραφή του σε ειδικά ερώτηση στο τέλος κάθε ερωτηματολογίου. Από την άλλη πλευρά ο υπολογισμός της Επιβάρυνσης Ανταπόκρισης σε οικονομικούς όρους γίνεται συνήθως με τον πολλαπλασιασμό του αριθμού των συμπληρωμένων ερωτηματολογίων με το μέσο κόστος ανά ώρα επί του χρόνου που διατέθηκε για τη συμπλήρωση του ερωτηματολογίου.

Σύμφωνα με την παραπάνω περιγραφή τα στοιχεία που θεωρείται σημαντικό να καταγράφονται στο τμήμα αυτό της ποιοτικής έκθεσης είναι:

9. Επιδόσεις, Κόστος και Επιβάρυνση Ανταπόκρισης

Κόστος

- ✓ Το επιχειρησιακό κόστος των καταγεγραμμένων βασικότερων κόστων για κάθε χρόνο.
- ✓ Πρόσφατες προσπάθειες που έγιναν για βελτίωση της αποδοτικότητας της στατιστικής διαδικασίας.
- ✓ Οι διαδικασίες που έγιναν για τον υπολογισμό της αποδοτικότητας.
- ✓ Η έκταση στην οποία διαδικασίες όπως η κωδικοποίηση δεδομένων, η επικύρωση και ο καταλογισμός αυτοματοποιήθηκαν.

Επιβάρυνση Ανταπόκρισης

- ✓ Η Επιβάρυνση Ανταπόκρισης εκφρασμένη σε χρόνο ή οικονομικούς όρους για κάθε χρόνο.
- ✓ Οι στόχοι μείωσης την Επιβάρυνσης Ανταπόκρισης.
- ✓ Πρόσφατες προσπάθειες που έγιναν για τη μείωση της.
- ✓ Κατά πόσο το εύρος και οι λεπτομέρειες των δεδομένων που συλλέχθηκαν μπορούν να περιοριστούν στις απολύτως απαραίτητες πληροφορίες.
- ✓ Κατά πόσο δεδομένα επιχειρήσεων είναι διαθέσιμα.
- ✓ Κατά πόσο συλλέγονται δεδομένα ηλεκτρονικά.

3.10. Εμπιστευτικότητα, Διαφάνεια και Ασφάλεια

Η Προστασία Προσωπικών Δεδομένων των προμηθευτών δεδομένων σε μία στατιστική διαδικασία (νοικοκυριά, επιχειρήσεις, καθώς και άλλων ερωτώμενων), η εμπιστευτικότητα των πληροφοριών που παρέχονται και η χρήση τους αποκλειστικά για στατιστικούς σκοπούς πρέπει να είναι απόλυτα διασφαλισμένα.

Ειδικότερα, όσον αφορά την εμπιστευτικότητα, η προστασία του απόρρητου είναι αυστηρά καθορισμένη από το νόμο ενώ το προσωπικό μίας έρευνας δεσμεύεται από το νόμο για την απόλυτη εχεμύθεια του. Η ποιοτική έκθεση είναι αναγκαίο να αναφέρεται στις συμφωνίες αυτές αλλά και σε τυχόν εξαιρέσεις από τις προαναφερθείσες νόμιμες διαδικασίες. Επιπροσθέτως, στο τμήμα αυτό της ποιοτικής έκθεσης πρέπει να περιλαμβάνεται η περιγραφή των διαδικασιών που διασφαλίζουν την εμπιστευτικότητα των αποτελεσμάτων κατά τη συλλογή, την επεξεργασία και τη μεταφορά των δεδομένων. Πρέπει να περιλαμβάνονται τα πρωτόκολλα σύμφωνα με τα οποία η πρόσβαση στα δεδομένα γίνεται μόνο για στατιστικούς σκοπούς, οι κανόνες σύμφωνα με τους οποίους αποκρύπτονται τα κατάλληλα πεδία από τους πίνακες που καταγράφονται τα δεδομένα και τέλος οι διαδικασίες με τις οποίες αποκλείεται οποιαδήποτε άλλη εναπομένουσα αποκάλυψη των δεδομένων που δίνονται εμπιστευτικά. Τέλος, θα πρέπει να αναφέρονται οι διαδικασίες που ισχύουν όταν γίνεται χρήση των δεδομένων για επιστημονικούς σκοπούς έτσι ώστε οι διαδικασίες αυτές να διασφαλίζουν την εμπιστευτικότητα των δεδομένων.

Επίσης, όσον αφορά τη διαφάνεια την έρευνας οι ερωτώμενοι θα πρέπει να ενημερώνονται για την εντολή με την οποία μία έρευνα συντάσσεται, τις χρήσεις των δεδομένων που συλλέγονται και τις συνθήκες που αφορούν την εμπιστευτικότητα τους. Επιπλέον οι ανακοινώσεις έναρξης της στατιστικής διαδικασίας καθώς και οι αναφορές στον τύπο επιβάλλεται να είναι αντικειμενικές και να μην εμπεριέχουν πολιτικά στοιχεία. Με σκοπό τη διατήρηση της διαφάνειας μίας στατιστικής διαδικασίας οι χρήστες είναι σημαντικό να ενημερώνονται για τους περιορισμούς που συνοδεύουν ως αναφορά τη χρήση τους τα δεδομένα. Οι πληροφορίες που συνοδεύουν τα δεδομένα είναι σημαντικό να δημοσιεύονται έτσι ώστε οι χρήστες των αποτελεσμάτων να μπορούν να καθορίσουν τη χρήση τους σε προσωπικό επίπεδο. Τέλος, σε περίπτωση εντοπισμού λαθών ή αποκλίσεων στα δημοσιευμένα δεδομένα είναι σημαντική η αναδημοσίευση τους καθώς με σκοπό τη διατήρηση της διαφάνειας. Η ποιοτική έκθεση στο τμήμα της Εμπιστευτικότητας, Διαφάνειας και Ασφάλειας κρίνεται απαραίτητο να περιγράφει όλες τις ρυθμίσεις που γίνονται με σκοπό τη διασφάλιση της διαφάνειας όλων των παραπάνω, καθώς επίσης είναι σημαντικό να επιδεικνύει τυχόν ελαττώματα που αφορούν τα ζητήματα αυτά.

Όσον αφορά την ασφάλεια των δεδομένων, στην ποιοτική έκθεση είναι αναγκαίο να περιέχονται οι διατάξεις οι οποίες διασφαλίζουν την ασφάλεια των διαδικασιών απόκτησης των δεδομένων, κυρίως όταν η διαδικασία αυτή γίνεται μέσω του Διαδικτύου, την ασφάλεια της αποθήκευσης των συμπληρωμένων εντύπων της έρευνας, δηλαδή πόσο καιρό διαρκεί η αποθήκευση καθώς και τις ασφαλείς διαδικασίες με τις οποίες απορρίπτονται, την ασφάλεια κρυπτογράφησης των δεδομένων όταν βρίσκονται σε υπολογιστές ή άλλες συσκευές, την ασφάλεια των συλλεχθέντων δεδομένων που βρίσκονται σε βάσεις δεδομένων καθώς και την ασφάλεια των δεδομένων που καταγράφονται αθροιστικά σε CD ή άλλες βάσεις δεδομένων. Τέλος, υπάρχουν διακανονισμοί που καθορίζουν τη σχέση μεταξύ της πρόσβασης στα δεδομένα και τη διατήρηση της εμπιστευτικότητας και της ασφάλειας τους. Οι διακανονισμοί αυτοί είναι επίσης σημαντικό να περιγράφονται στην ποιοτική έκθεση.

Συνεπώς, τελικά έχουμε ότι στο τμήμα της Εμπιστευτικότητας, Διαφάνειας και Ασφάλειας μίας ποιοτικής έκθεσης κρίνεται αναγκαίο να περιλαμβάνονται οι παρακάτω πληροφορίες:

10. Εμπιστευτικότητα, Διαφάνεια και Ασφάλεια

- ✓ Καταγραφή αν και εφόσον η εμπιστευτικότητα απαιτείται από το νόμο και οι συμφωνίες εμπιστευτικότητας του προσωπικού μιας στατιστικής διαδικασίας.
- ✓ Κατά πόσο επιστημονικές έρευνες μπορούν να έχουν πρόσβαση στα δεδομένα και εάν ναι οι συνθήκες με τις οποίες διασφαλίζεται η εμπιστευτικότητα.
- ✓ Οι διαδικασίες με τις οποίες διασφαλίζεται η εμπιστευτικότητα των αποτελεσμάτων καθόλη τη διάρκεια μίας στατιστικής διαδικασίας.
- ✓ Ενημέρωση των ερωτώμενων για τις χρήσεις των δεδομένων καθώς και των διαδικασιών που διασφαλίζουν την εμπιστευτικότητα των δεδομένων.
- ✓ Κατά πόσο οι ανακοινώσεις έναρξης της στατιστικής διαδικασίας είναι αντικειμενικές.
- ✓ Κατά πόσο οι αποκλίσεις που εντοπίζονται στα δεδομένα οδηγούν σε αναδημοσίευση τους.
- ✓ Αναφορά των διατάξεων σύμφωνα με τις οποίες διασφαλίζεται η ασφάλεια των δεδομένων.

3.11. Συμπεράσματα

Το τελευταίο τμήμα της ποιοτικής έκθεσης κρίθηκε σημαντικό να περιλαμβάνει αναφορά στα κυριότερα ποιοτικά θέματα που απασχόλησαν τη στατιστική διαδικασία και κατ' επέκταση την ποιοτική έκθεση καθώς επίσης και συστάσεις που στοχεύετε να βελτιώσουν ή να επιλύσουν τα ζητήματα αυτά. Τέλος, είναι εξίσου σημαντικό να αναφέρονται οι χρήστες και οι χρήσεις των αποτελεσμάτων της στατιστικής διαδικασίας αλλά και οποιαδήποτε διαδικασία αναμένεται να ακολουθήσει.

Συνεπώς έχουμε ότι στο τελευταίο τμήμα της ποιοτικής έκθεσης αναφέρονται πληροφορίες που αφορούν:

11. Συμπεράσματα

- ✓ Αναφορά των κυριότερων ποιοτικών ζητημάτων της στατιστικής διαδικασίας.
- ✓ Συστάσεις που στοχεύουν στη βελτίωση ή επίλυση των ζητημάτων που προέκυψαν.
- ✓ Αναφορά των διαδικασιών που πρόκειται να ακολουθήσουν.

Κεφάλαιο 4

Πληροφορίες τμημάτων ποιοτικής έκθεσης της έρευνας EU-SILC

Οι παρακάτω πληροφορίες αφορούν τις συγκεντρωτικές πληροφορίες που προέκυψαν από τις ποιοτικές εκθέσεις των χωρών για την EU-SILC. Οι πληροφορίες δημοσιεύονται από την Eurostat και αφορούν τμήματα μίας ποιοτικής έκθεσης. Φυσικά, η πρόσβαση στις ποιοτικές εκθέσεις των χωρών δεν είναι δημόσια.

4.1. Εισαγωγή στη στατιστική διαδικασία και στα αποτελέσματά της.

4.1.1. Η στατιστική διαδικασία

Από το 1994 – 2001 η βασική πηγή δεδομένων για τους τομείς που καλύπτει η EU-SILC ήταν όπως έχει προαναφερθεί η ECHP (European Community Household Panel). Η περίοδος 2001-2005 αποτέλεσε μία μεταβατική περίοδο όπου εθνικά δεδομένα εναρμονιζόντουσαν έτσι ώστε να αποτελέσουν πηγή για τη μελέτη των δεικτών που μελετά η EU-SILC. Στις περισσότερες περιπτώσεις οι χώρες ξεκίνησαν την EU-SILC από το μηδέν με τη συλλογή ολοκληρωμένων σύγχρονων και διαχρονικών δεδομένων. Άλλες χώρες χρησιμοποίησαν ένα συνδυασμό από συνεντεύξεις και χρήση των εθνικών μητρώων τους. Τέλος υπάρχουν χώρες που επιδιώκουν να προσαρμόσουν τις ήδη υπάρχουσες εθνικές πηγές.

Το ελάχιστο αποτελεσματικό μέγεθος του μεγέθους του δείγματος για τα σύγχρονα δεδομένα καθορίζεται από τον κανονισμό της EU-SILC και βέβαια έχει εκτενώς περιγραφεί στο κεφάλαιο 1. Σύμφωνα με τα στοιχεία που ισχύουν έως τώρα όσον αφορά το μέγεθος του ελάχιστου αποτελεσματικού δείγματος ανέρχεται στα 130.000 νοικοκυριά και 273.000 άτομα και συμφωνεί με την εκτενή περιγραφή που έχει γίνει.

Η συλλογή των δεδομένων για την έρευνα EU-SILC γίνεται κάθε χρόνο ενώ όσον αφορά τη συλλογή τους γίνεται είτε από τα μητρώα των χωρών είτε από συνεντεύξεις. Οι συνεντεύξεις μπορούν να γίνουν με 4 διαφορετικές μεθόδους:

1. Με προσωπική συνέντευξη η οποία γίνεται με μολύβι και χαρτί (PAPI-Paper Assisted Personal Interview).
2. Με προσωπική συνέντευξη και τη χρήση ηλεκτρονικού υπολογιστή (CAPI-Computer-Assisted Personal Interview).
3. Με τηλεφωνική συνέντευξη και τη χρήση ηλεκτρονικού υπολογιστή (CATI-Computer-Assisted Telephone Interview).
4. Αυτοδιαχειριζόμενο ερωτηματολόγιο.

Οι συνολικές εκτιμήσεις προκύπτουν από τον υπολογισμό του μέσου όρου των ανεξάρτητων εθνικών εικόνων με βαρύτητα ανάλογη του πληθυσμού της κάθε χώρας και τέλος ο καταλογισμός πεδίων γίνεται σύμφωνα με τον κανονισμό της EU-SILC που αφορά το τμήμα του καταλογισμού πεδίων.

4.1.2. Περιγραφή των συλλεγόμενων δεδομένων

Η ευρωπαϊκή έρευνα EU-SILC που αφορά το εισόδημα και τις συνθήκες ζωής καλύπτει 4 βασικά θέματα. Με την έννοια κάλυψη εννοούμε τη συλλογή πληροφοριών για τη μελέτη των 4 βασικών θεμάτων μέσα από τους κατάλληλους δείκτες. Το πρώτο ζήτημα αφορά τους βασικούς δείκτες της έρευνας- οι οποίοι θα διευκρινιστούν παρακάτω, το δεύτερο ζήτημα αφορά τη μελέτη της κατανομής του εισοδήματος και της «νομισματικής» φτώχειας, το τρίτο ζήτημα τις συνθήκες ζωής και το τέταρτο ζήτημα αφορά την υλική στέρηση και τους δείκτες που αφορούν τη περιγραφή της ρύθμισης της παιδικής φροντίδας. Η δομή των ζητημάτων καλύπτεται από το σύνολο των δεικτών που μελετώνται.

Οι βασικοί δείκτες της EU-SILC αναφέρονται στους δείκτες που παρέχονται από το πλαίσιο του OMC (σχετικά καινούργιου, διακυβερνητικού μέσου στην Ευρωπαϊκή Ένωση του επονομαζόμενου “Open Method of Coordination”) και αφορούν την καταπολέμηση της φτώχειας και του κοινωνικού αποκλεισμού. Το OMC βασίζεται στην εθελοντική συνεργασία των κρατών μελών της Ευρωπαϊκής Ένωσης και ένα από τα αντικείμενα του είναι ο καθορισμός ποσοτικών και ποιοτικών δεικτών σύμφωνα με τις ανάγκες των κρατών μελών της Ευρωπαϊκής Ένωσης με σκοπό την εξασφάλιση της καλύτερης δυνατής συγκρισιμότητας. Οι βασικοί δείκτες της EU-SILC πιο συγκεκριμένα αποτελούνται από 3 ομάδες δεικτών. Η πρώτη ομάδα αφορά τους δείκτες για την κοινωνική ένταξη, η δεύτερη ομάδα αφορά τις συντάξεις των ερωτώμενων και η τρίτη ομάδα αφορά τους πρωταρχικούς δείκτες οι οποίοι είναι μία λίστα από 14 δείκτες κοινοί για όλες τις χώρες, οι οποίοι εξασφαλίζουν τη μέγιστη συγκρισιμότητα και τον υπολογισμό της προόδου όσον αφορά το επίπεδο της τιμής τους.

Το δεύτερο ζήτημα μελέτης της κατανομής του εισοδήματος και της φτώχειας αφορά δείκτες οι οποίοι σχετίζονται με τον κίνδυνο της φτώχειας, τον κίνδυνο φτώχειας των εργαζόμενων ανεξάρτητων ατόμων, το εισόδημα των ατόμων που βρίσκονται υπό τον κίνδυνο της φτώχειας αλλά και δείκτες που αφορούν τη μελέτη της κατανομής του εισοδήματος.

Όσον αφορά τη μελέτη των συνθηκών ζωής από την EU-SILC υπάρχουν δείκτες οι οποίοι αφορούν τα χαρακτηριστικά και τις συνθήκες ζωής των νοικοκυριών, τα χαρακτηριστικά του πληθυσμού ανάλογα με τις διασπάσεις του σε υποκατηγορίες, την υγεία, τις συνθήκες εργασίας και διαμονής στο νοικοκυριό.

Οι δείκτες που αφορούν τη μελέτη της υλικής στέρησης σχετίζονται με την διαρκή οικονομική στενότητα, τη στέγαση και το περιβάλλον της κατοικίας. Η μελέτη των δεικτών που αφορούν τις ρυθμίσεις φροντίδας των παιδιών αφορά δείκτες που σχετίζονται με τη φροντίδα των παιδιών.

4.2. Συνάφεια (Relevance)

Οι βασικοί χρήστες της EU-SILC και των αποτελεσμάτων της είναι ιδρύματα όπως υπηρεσίες της Ευρωπαϊκής Ένωσης, Ευρωπαϊκά Ινστιτούτα, Εθνικές Διοικητικές Μονάδες (κυρίως εκείνες που ασχολούνται με τη προστασία των κατοίκων, την κοινωνική ένταξη) ή άλλοι διεθνείς οργανισμοί, χρήστες που ασχολούνται με τη στατιστική στη Eurostat ή στις Εθνικές Στατιστικές Αρχές των χωρών με σκοπό τη δημοσίευση αποτελεσμάτων. Επίσης βασικούς χρήστες αποτελούν και οι ερευνητές που έχουν πρόσβαση στα συλλεχθέντα δεδομένα αλλά και Μέσα Μαζικής Ενημέρωσης που ενδιαφέρονται για τις συνθήκες ζωής και την κοινωνική συνοχή των κατοίκων της Ευρωπαϊκής Ένωσης.

Η γεωγραφική κάλυψη της EU-SILC αναφέρεται εκτενώς στο Κεφάλαιο 1. Η EU-SILC συμπεριλαμβάνει άτομα που μένουν σε νοικοκυριά ηλικίας αυστηρά άνω των 16 ετών, όπως έχει προαναφερθεί, ενώ άτομα που μένουν σε συλλογικά καταλύματα και ινστιτούτα δεν περιλαμβάνονται στον πληθυσμό. Επίσης, γεωγραφικά τμήματα τα οποία δεν περιλαμβάνουν ποσοστό κατοίκων μεγαλύτερο του 2% και συνήθως είναι δυσπρόσιτα συχνά δεν περιλαμβάνονται και το γεγονός αυτό αναφέρεται από τις χώρες.

4.3. Ακρίβεια

Σύμφωνα με τον κανονισμό της EU-SILC που αφορά τη δειγματοληψία και τους κανόνες της για όλα τα μέλη της EU-SILC ισχύει ότι τόσο τα σύγχρονα όσο και τα διαχρονικά δεδομένα βασίζονται σε ένα αντιπροσωπευτικό δείγμα ίσης πιθανότητας επιλογής νοικοκυριών της κάθε χώρας, ανεξάρτητου της γλώσσας και εθνικότητας. Όπως έχει προαναφερθεί το πλαίσιο του δείγματος και οι μέθοδοι επιλογής πρέπει να διασφαλίζουν την ίση και μη μηδενική πιθανότητα επιλογής για όλα τα νοικοκυριά και τα άτομα.

Όσον αφορά τα σφάλματα δειγματοληψίας αυτά χρησιμοποιούνται για τον υπολογισμό της αξιοπιστίας των δεδομένων που συλλέγονται σε μία έρευνα που βασίζεται στη δειγματοληψία. Η EU-SILC σχεδιάστηκε για να υπολογίσει το ποσοστό κατοίκων που βρίσκονται υπό τον κίνδυνο της φτώχειας με ακρίβεια της τάξεως της μονάδας του εκατοστού. Τα μεγέθη των δειγμάτων καθορίζονται σύμφωνα με τη συγκεκριμένη απαίτηση ακρίβειας.

Σφάλματα Κάλυψης

Δύο βασικές πηγές σφαλμάτων κάλυψης εντοπίστηκαν:

- Οι χώρες που βασίστηκαν σε πληροφορίες για τα νοικοκυριά από τα εθνικά μητρώα διότι για να επιτευχθεί η βέλτιστη κάλυψη πρέπει τα μητρώα να ανανεώνονται με την παραμικρή μετακίνηση του πληθυσμού (άτομα που περιλαμβάνονται σε ένα νοικοκυριό ή το εγκαταλείπουν).

- Οι χώρες που χρησιμοποίησαν εθνικές βάσεις δεδομένων στις οποίες περιλαμβάνονται το σύνολο των διευθύνσεων. Και σε αυτή την περίπτωση πρέπει να διασφαλίζεται η συνεχής ανανέωση της βάσης.
- Ο χρόνος που υπάρχει μεταξύ της επιλογής δείγματος και της περιόδου που συλλέγονται οι πληροφορίες από το δείγμα. Ο χρόνος αυτός πρέπει να είναι όσο το δυνατό μικρότερος.
- Επίσης, στην περίπτωση που οι χώρες επιλέγουν το δείγμα τους σύμφωνα με την ανταπόκριση των «διευθύνσεων» σε άλλες έρευνες μπορεί να οδηγήσει σε σφάλμα κάλυψης.

Σφάλματα Μέτρησης και Επεξεργασίας

Λόγω της πολυπλοκότητας και της ευαισθησίας της έρευνας, οι ερωτώμενοι δεν μπορούσαν ή δε θέλανε να δώσουν πληροφορίες για όλα τα εισοδήματα τους. Για παράδειγμα σε περίπτωση προσωπικής εργασίας η ανταμοιβή δεν αναφέρθηκε από τους ερωτώμενους. Επίσης η EU-SILC ενδιαφέρεται για το εισόδημα π.χ. που προκύπτει από την προσωπική χρήση ενός αυτοκινήτου που ανήκει σε εταιρεία, δηλαδή το μη «νομισματικό εισόδημα». Συνεπώς, ο κίνδυνος σύγχυσης των πληροφοριών που πρέπει να δοθούν αυξάνεται.

Σφάλματα μη ανταπόκρισης

Η έρευνα EU-SILC υποφέρει τόσο από τη μη ανταπόκριση μονάδων του πληθυσμού (unit non-response) όσο και από τη μη ανταπόκριση σε συγκεκριμένες ερωτήσεις από κάθε μονάδα του πληθυσμού (item non-response) συνήθως λόγω του ότι θεωρούνται προσωπικές ή εμπιστευτικές από τον ερωτώμενο. Η μη ανταπόκριση σε συγκεκριμένες ερωτήσεις είναι αυξημένη όταν το περιεχόμενό τους αφορά το εισόδημα και συνήθως ξεπερνάτε με τη μέθοδο του καταλογισμού. Η μέθοδος του καταλογισμού δεν αφορά αληθινές τιμές των πεδίων που καταλογίστηκαν με αποτέλεσμα να έχει σημαντική επίδραση στη συνολική ακρίβεια.

Το ανώτατο ποσοστό μη ανταπόκρισης των επιλεγμένων νοικοκυριών/ατόμων δε πρέπει να ξεπερνά το 40% ενώ η συνολική μη ανταπόκριση σε ερωτήσεις που αφορούν το εισόδημα πρέπει να περιορίζεται στο 5%.

4.4. Επικαιρότητα και Εγκαιρότητα (Timeliness and Punctuality)

Τα σύγχρονα δεδομένα είναι διαθέσιμα σε πίνακες 12 μήνες μετά το τέλος της συλλογής δεδομένων ενώ τα διαχρονικά δεδομένα είναι διαθέσιμα 18 μήνες μετά τη διαδικασία συλλογής τους. Επιπρόσθετα, ανώνυμοι φάκελοι που περιέχουν σύγχρονα δεδομένα για ερευνητικούς σκοπούς είναι διαθέσιμα 15 μήνες μετά το τέλος της συλλογής τους ενώ αντίστοιχοι φάκελοι με διαχρονικά δεδομένα 20 μήνες μετά το τέλος της συλλογής τους.

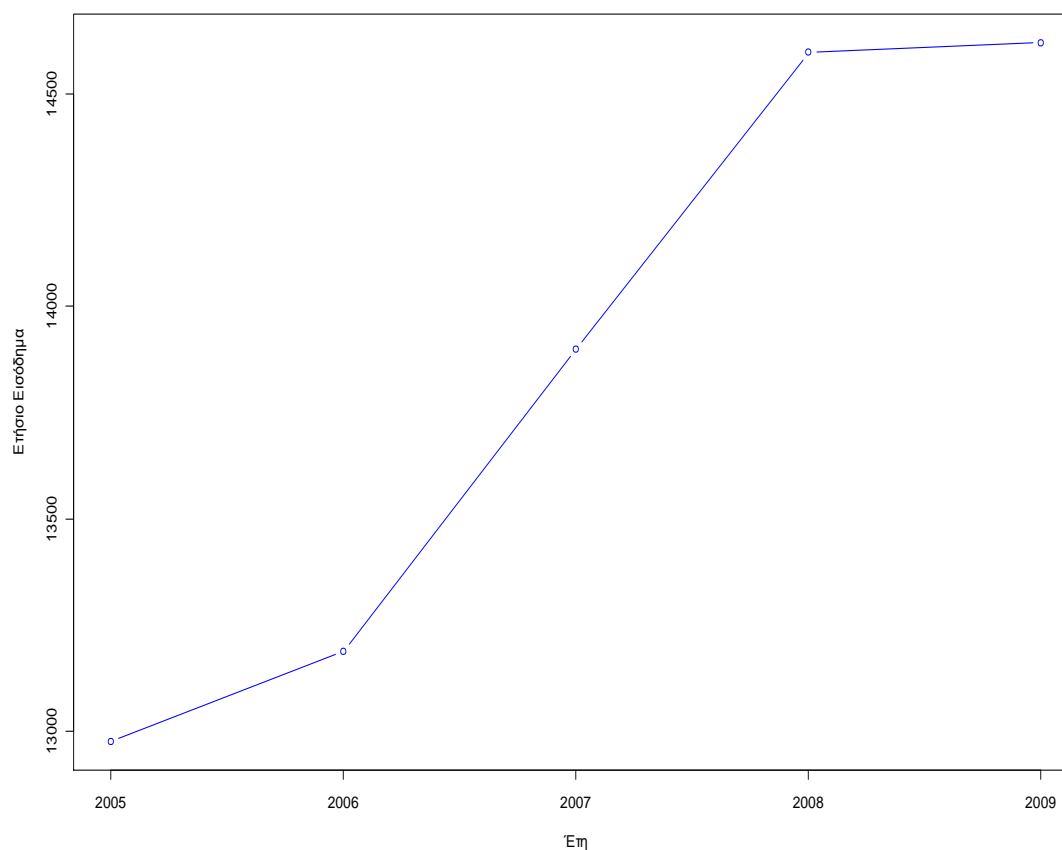
4.5. Συγκρισιμότητα (Comparability)

Όπως αναφέρθηκε η *συγκρισιμότητα* των αποτελεσμάτων είναι μία ειδική περίπτωση της συνοχής των αποτελεσμάτων και πιο συγκεκριμένα η περίπτωση όπου τα αποτελέσματα της έρευνας αναφέρονται στα ίδια στατιστικά δεδομένα αλλά στοχεύουν σε χρονική σύγκριση, σύγκριση μεταξύ περιοχών ή άλλων συγκρίσεων.

4.5.1 Ετήσιο Εισόδημα

Στην έρευνα EU-SILC, είναι γνωστό ότι ένας από τους σημαντικότερους δείκτες που στοχεύεται να εκτιμηθεί είναι η μέση τιμή του εισοδήματος (€) για κάθε χώρα που συμμετέχει στην έρευνα. Παρακάτω δίνεται η χρονοσειρά της μέσης τιμής του ετήσιου εισοδήματος των 27 κρατών μελών της Ευρωπαϊκής Ένωσης για τις χρονιές 2005-2009.

Το ετήσιο εισόδημα (€) των 27 κρατών μελών της Ευρωπαϊκής Ένωσης το διάστημα 2005-2009.



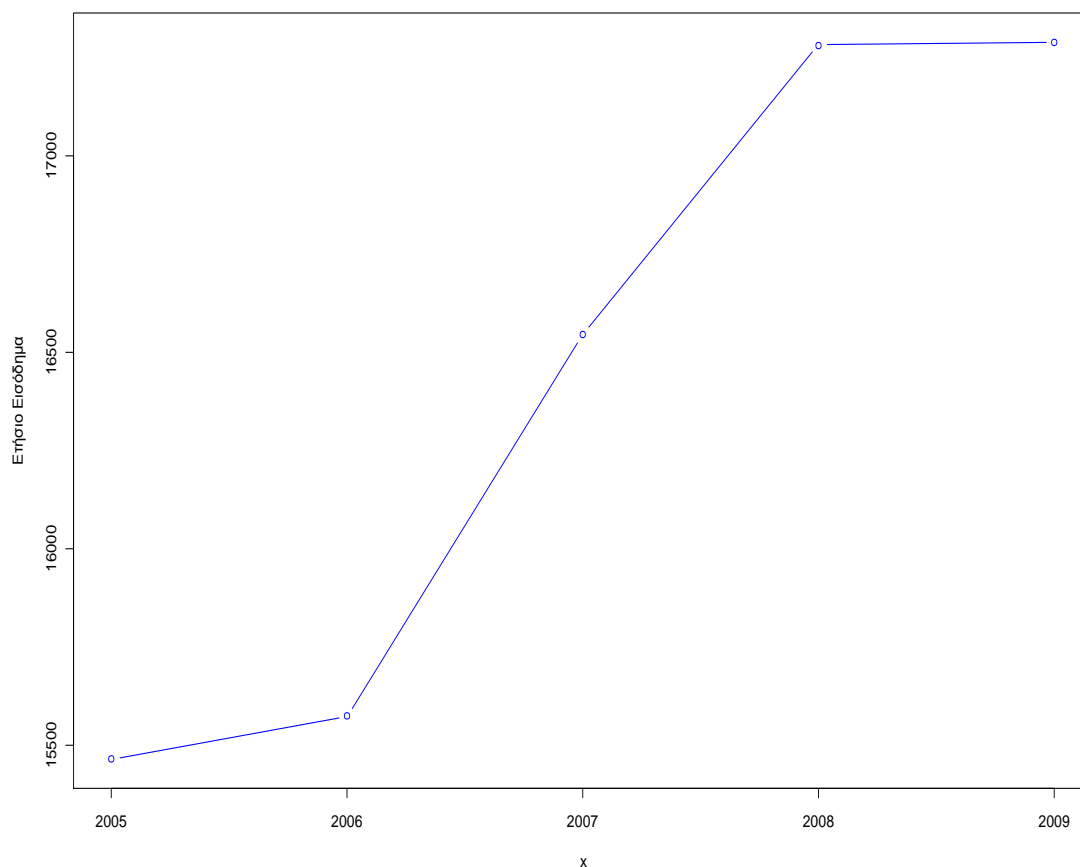
Οι τιμές αφορούν τη μέση τιμή του ετήσιου εισοδήματος που υπολογίστηκε μέσα από την έρευνα για τα έτη 2005 έως 2009. Σύμφωνα με τις τιμές της χρονοσειράς είναι φανερό ότι υπάρχει συγκρισιμότητα με την πάροδο του χρόνου για τη μεταβλητή του εισοδήματος. Οι τιμές των εισοδημάτων δεν έχουν μεγάλες διαφορές από έτος σε έτος. Πιο συγκεκριμένα, αν συμβολίσουμε με I_n τη μέση τιμή του εισοδήματος για το έτος n που δημοσιεύθηκε για τις 27 χώρες της Ευρωπαϊκής

Ένωσης έχουμε τη παρακάτω σχέση: $I_n \leq I_{n-1} + 800 \text{ €}$ η οποία ισχύει για όλα τα έτη. (2005-2009).

Η τάξη μεγέθους των 800€ αποδεικνύει ότι υπάρχει συγκρισιμότητα με την πάροδο του χρόνου για τη μεταβλητή της μέσης τιμής του εισοδήματος των 27 κρατών μελών της Ευρωπαϊκής Ένωσης. Η μελέτη των αλλαγών των οικονομικών καταστάσεων των χωρών δικαιολογεί πλήρως το εύρος του μεγέθους. Συνεπώς, δε τίθεται ζήτημα αλλαγής της στατιστικής διαδικασίας λόγω διαφορετικών μεθόδων που χρησιμοποιήθηκαν ή διαφορετικού σκεπτικού της έρευνας EU-SILC από έτος σε έτος, με αποτέλεσμα να διαφαίνεται η συγκρισιμότητα στη μεταβλητή του μέσου εισοδήματος.

Με ανάλογη διαδικασία έγινε φανερή η συγκρισιμότητα με την πάροδο του χρόνου των αποτελεσμάτων που αφορούν τη μέση τιμή του ετήσιου εισοδήματος των 15 κρατών μελών της Ευρωπαϊκής Ένωσης. Στο τέλος της παρούσας εργασίας δίνονται οι πίνακες των 27 κρατών μελών και των 15 κρατών μελών της Ευρωπαϊκής Ένωσης.

Το ετήσιο εισόδημα (€) των 15 κρατών μελών της Ευρωπαϊκής Ένωσης το διάστημα 2005-2009.



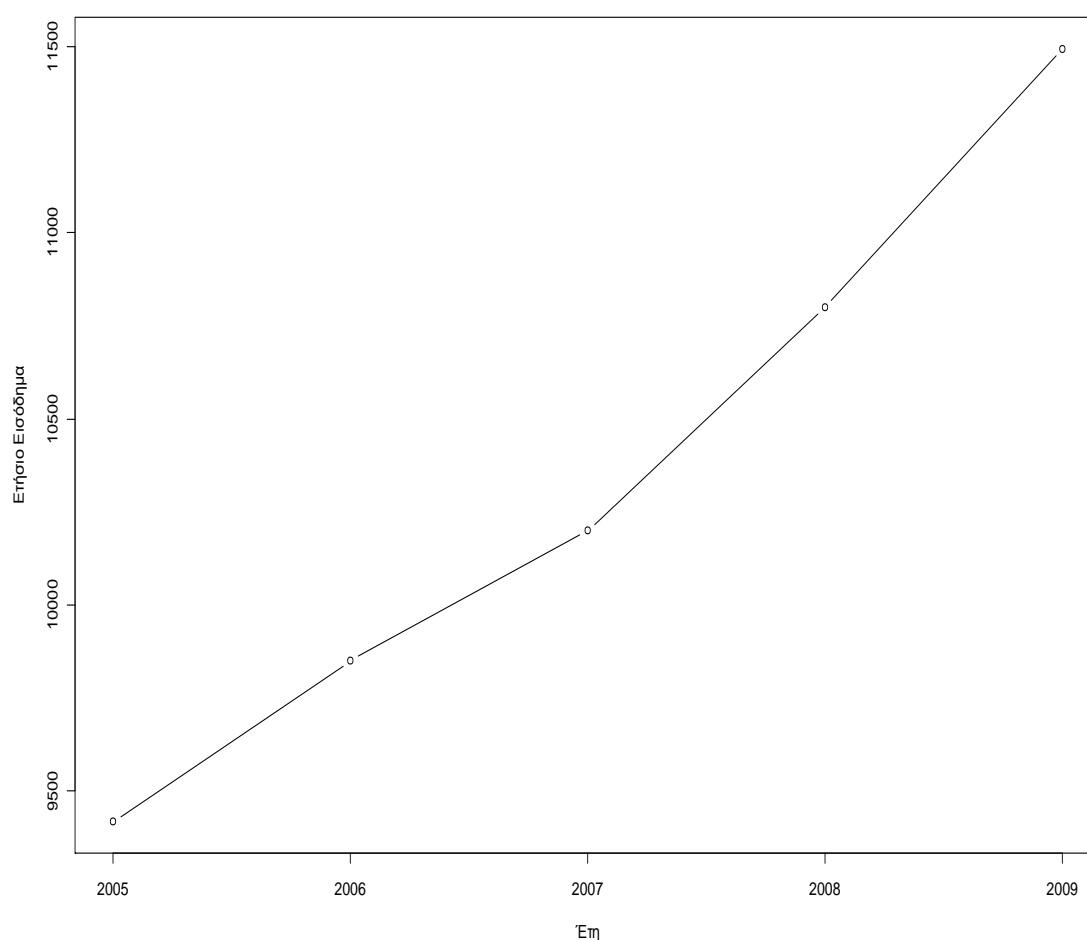
Παρατηρείται φυσικά ότι η μεγαλύτερη αύξηση του εισοδήματος έγινε από το 2006 στο 2007 και από το 2007 στο 2008, όπως ακριβώς και στη περίπτωση του μέσου εισοδήματος των 27 κρατών μελών της Ευρωπαϊκής Ένωσης. Επιπλέον, και στην

περίπτωση του εισοδήματος των 15 κρατών μελών της Ευρωπαϊκής Ένωσης υπάρχει συγκρισιμότητα με την πάροδο του χρόνου. Πιο συγκεκριμένα προκύπτει η ανίσωση: $I_n \leq I_{n-1} + 970\text{€}$ η οποία ισχύει για όλα τα έτη.

Η σχέση αποδεικνύει και σε αυτή την περίπτωση τη δυνατότητα συγκρισιμότητας με την πάροδο του χρόνου για το ετήσιο εισόδημα των 15 κρατών μελών της Ευρωπαϊκής Ένωσης.

Η διαδικασία επαναλήφθηκε για όλες τις χώρες που συμμετέχουν στην EU-SILC και τα αποτελέσματα ήταν ανάλογα με τις δύο παραπάνω μελέτες. Παρουσιάζεται επίσης η περίπτωση της μέσης τιμής του ετήσιου εισοδήματος για την Ελλάδα για την ίδια χρονική περίοδο (2005 έως και 2009).

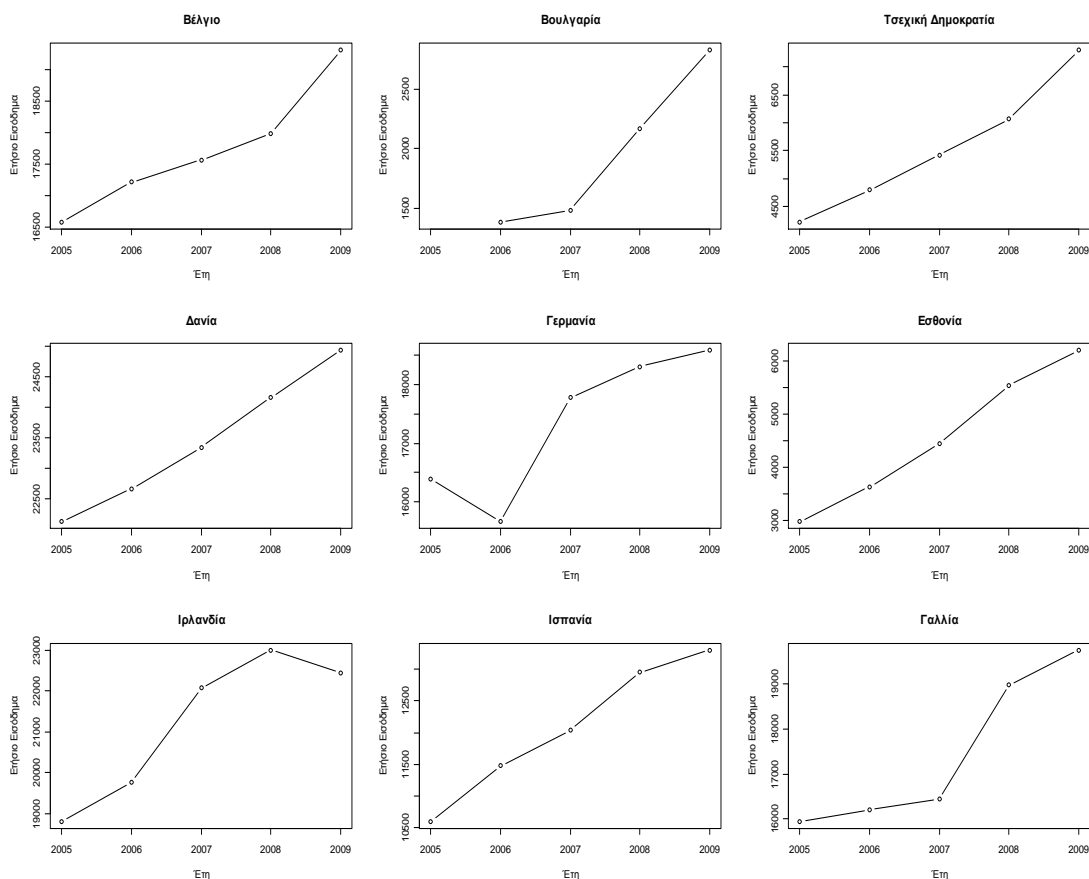
Το ετήσιο εισόδημα της Ελλάδος το διάστημα 2005-2009.

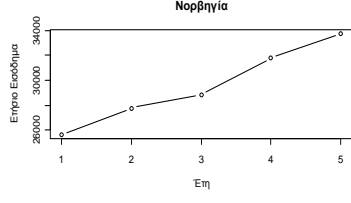
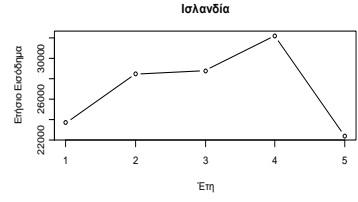
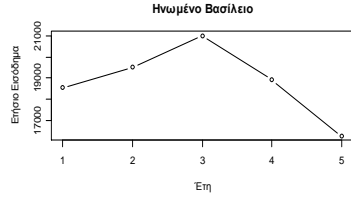
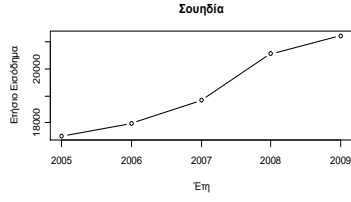
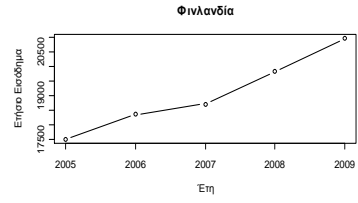
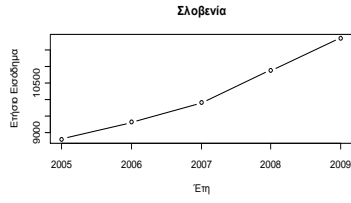
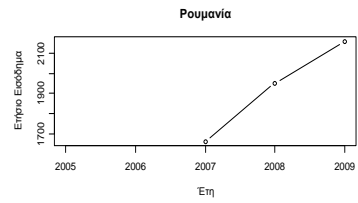
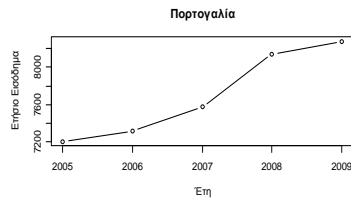
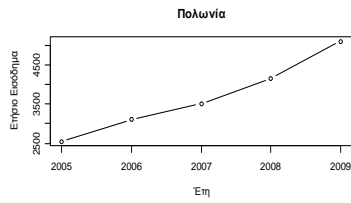
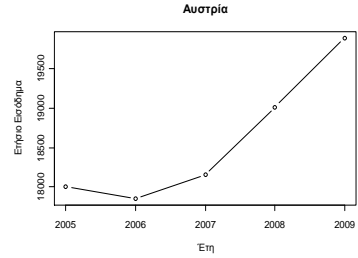
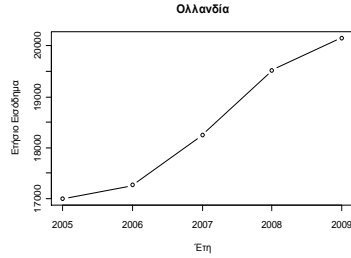
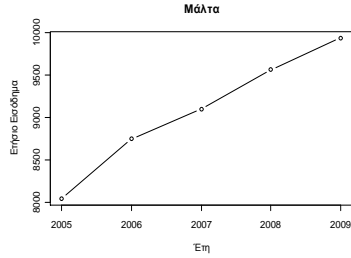
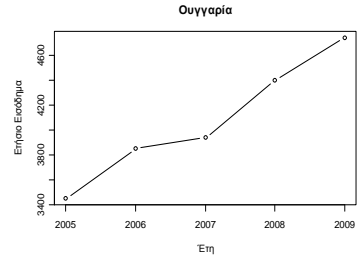
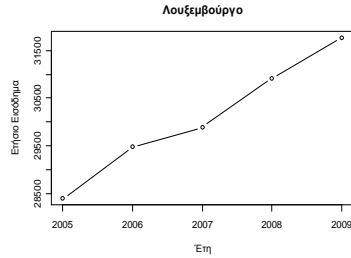
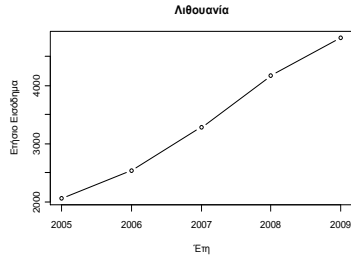
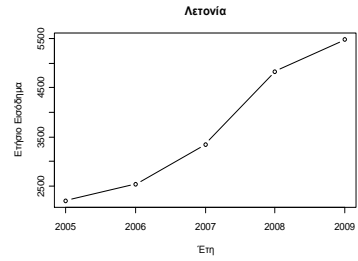
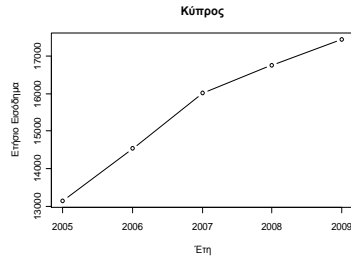
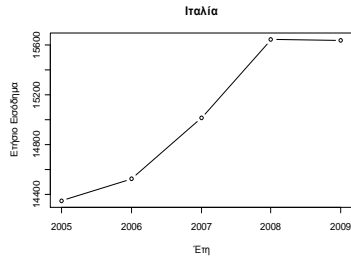


Για την περίπτωση της χώρας της Ελλάδος το ετήσιο εισόδημα της χρονιάς n αποκλίνει από το ετήσιο εισόδημα του προηγούμενου έτους το πολύ κατά 700€. Η τάξη μεγέθους και στην περίπτωση της μίας χώρας (Ελλάδα) αποδεικνύει τη δυνατότητα συγκρισιμότητας με την πάροδο του χρόνου. Πιο συγκεκριμένα, για όλα τα έτη:

$$I_n \leq I_{n-1} + 700\text{€}.$$

Τέλος, δίνονται παρακάτω οι γραφικές παραστάσεις του ετήσιου εισοδήματος των χωρών που συμμετέχουν στην έρευνα. Χωρίς περαιτέρω ανάλυση είναι φανερό ότι οι τιμές όλων των χωρών από έτος σε έτος είναι συγκρίσιμες με την πάροδο του χρόνου και επίσης βλέπουμε ότι η «εικόνα» της τιμής του ετήσιου εισοδήματος για κάθε χώρα έρχεται σε συμφωνία με την «εικόνα» τόσο της τιμής του ετήσιου εισοδήματος των 27 κρατών μελών όσο και της τιμής του ετήσιου εισοδήματος των 15 κρατών μελών της Ευρωπαϊκής Ένωσης.





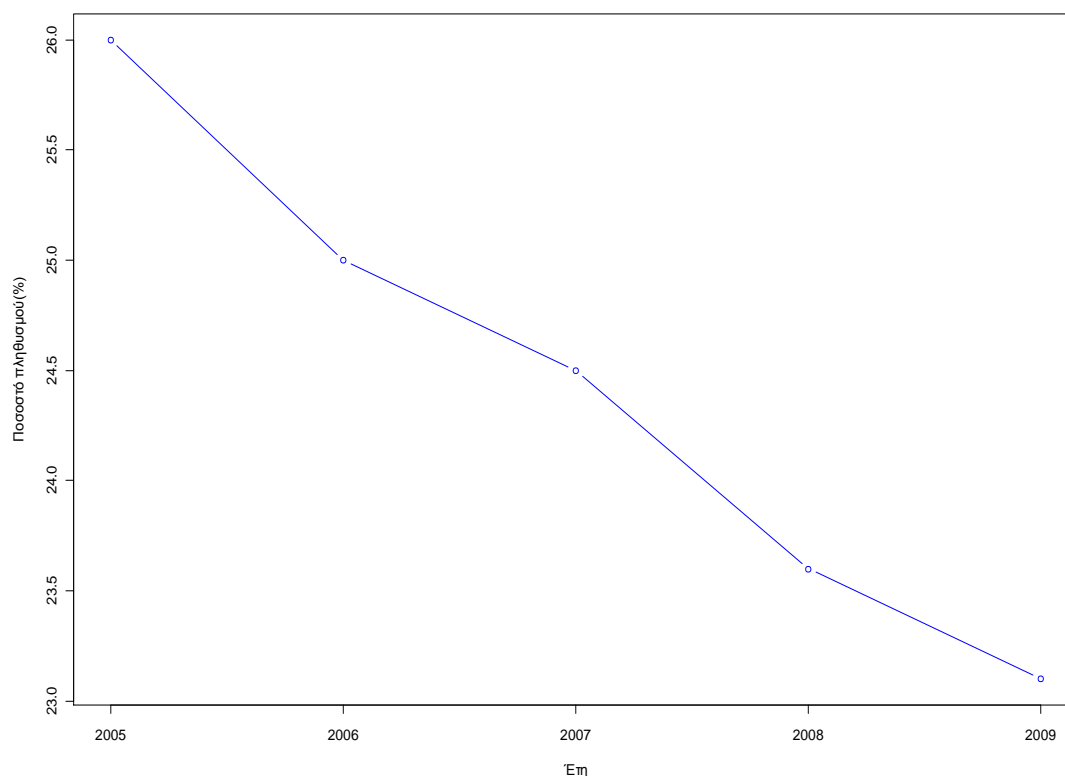
Όσον αφορά τις χώρες Ελβετία και Τουρκία έχει δημοσιευθεί μόνο μία τιμή ετήσιου εισοδήματος (δηλαδή η τιμή του εισοδήματος για μόνο ένα έτος από το 2005-2009) με αποτέλεσμα να μη μπορεί να γίνει έλεγχος συγκρισιμότητας με την πάροδο του χρόνου για τις δύο αυτές χώρες που συμμετέχουν στην έρευνα.

4.5.2 Ποσοστό φτώχειας

Μία δεύτερη σημαντική μεταβλητή είναι η μεταβλητή του ποσοστού φτώχειας. Το ποσοστό φτώχειας, έχει αναφερθεί και σε προηγούμενα κεφάλαια, καθορίζεται από το ποσοστό των ανθρώπων που παραμένουν στη φτώχεια (δηλαδή έχουν κατά 60% εισόδημα μικρότερο από τη τιμή της διαμέσου του εισοδήματος) για τουλάχιστον ένα αριθμό συνεχόμενων ετών. Στα αποτελέσματα της EU-SILC δημοσιεύεται το ποσοστό αυτό για κάθε χώρα που συμμετέχει στην έρευνα και αφορά φυσικά το ποσοστό του πληθυσμού κάθε χώρας που παραμένει στη φτώχεια. Με όμοιο τρόπο με εκείνον που χρησιμοποιήθηκε στο τμήμα του ετησίου εισοδήματος, θα ελεγχθεί η συγκρισιμότητα με την πάροδο του χρόνου για τη μεταβλητή του ποσοστού φτώχειας της έρευνας EU-SILC.

Σε πρώτο στάδιο, ξεκινάμε από τον έλεγχο της συγκρισιμότητας με την πάροδο του χρόνου μέσα από τις τιμές του ποσοστού του πληθυσμού των 27 κρατών μελών της Ευρωπαϊκής Ένωσης που δημοσιεύθηκαν για την χρονολογική περίοδο 2005-2009.

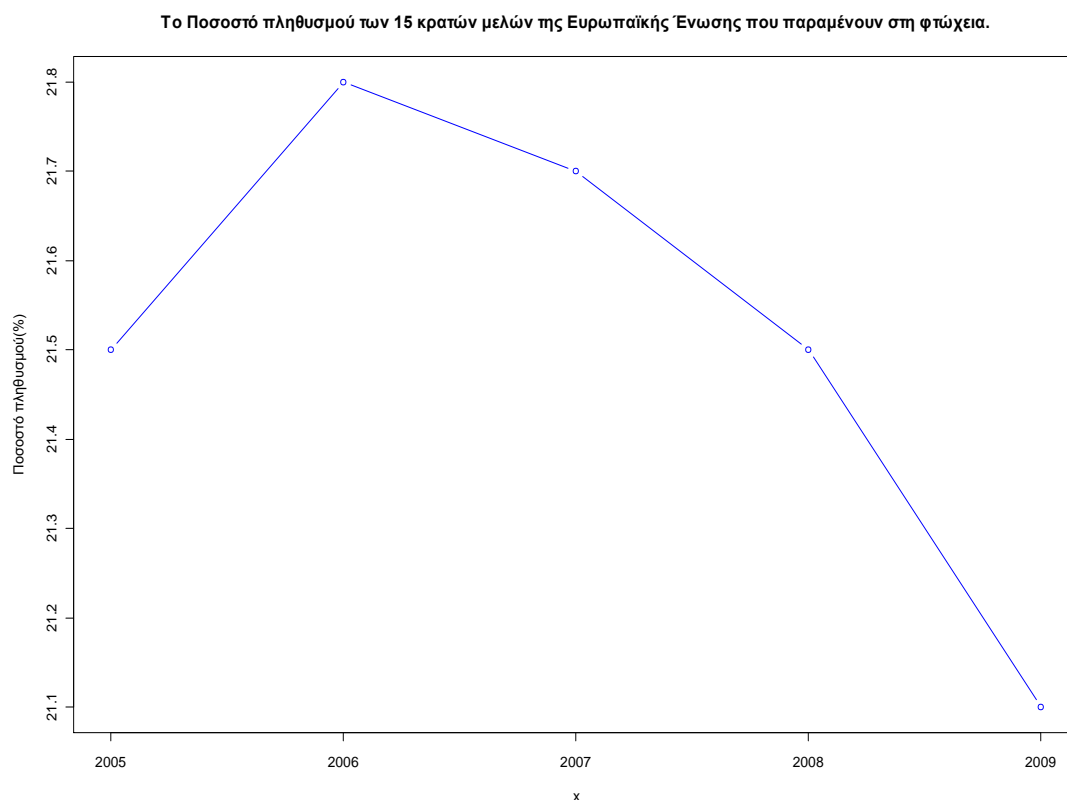
Το Ποσοστό πληθυσμού των 27 κρατών μελών της Ευρωπαϊκής Ένωσης που παραμένουν στη φτώχεια.



Οι τιμές του ποσοστού της φτώχειας κυμαίνονται από 23.1% έως 26%. Οι διακυμάνσεις που παρατηρούνται από έτος σε έτος δεν είναι τόσο μεγάλες ώστε να μην δικαιολογούνται από τις συνθήκες ζωής που άλλαξαν από έτος σε έτος και έρχονται σε πλήρη συμφωνία με τη μέση τιμή του εισοδήματος η οποία αυξάνεται. Επιπλέον, οι διαφορές στις τιμές των ποσοστών φτώχειας ανά έτος δεν είναι τόσο σημαντικές ώστε να υποτεθεί ότι η στατιστική διαδικασία που ακολουθήθηκε άλλαξε από το ένα έτος στο επόμενο. Πιο συγκεκριμένα παρατηρήθηκε ότι αν συμβολίσουμε με Π_n το ποσοστό του πληθυσμού για το έτος n προκύπτει: $\Pi_{n-1} \leq \Pi_n + 1\%$. Η σχέση επαληθεύεται για όλα τα έτη.

Η διαφορά του 1% του ποσοστού του πληθυσμού που παραμένουν στη φτώχεια αποδεικνύει ότι οι τιμές είναι συγκρίσιμες μεταξύ τους και συνεπώς και για τη μεταβλητή του ποσοστού του πληθυσμού στη φτώχεια των 27 κρατών μελών υπάρχει συγκρισιμότητα με την πάροδο του χρόνου.

Ομοίως, μελετάται ακολούθως η ύπαρξη συγκρισιμότητας με την πάροδο του χρόνου για το ποσοστό του πληθυσμού που δημοσιεύθηκε των 15 κρατών μελών της Ευρωπαϊκής Ένωσης.



Από τη γραφική παράσταση του ποσοστού φτώχειας για τον πληθυσμό των 15 κρατών μελών της Ευρωπαϊκής Ένωσης παρατηρούμε τόσο τη μείωση του ποσοστού το οποίο κυμαίνεται από 21.1% έως 21.8% όσο και την αύξηση του ποσοστού μεταξύ του έτους 2005 και του έτους 2006. Οι δύο αυτές καταστάσεις δικαιολογούνται από τη σύνθεση των 15 κρατών μελών και των 27 κρατών μελών της Ευρωπαϊκής

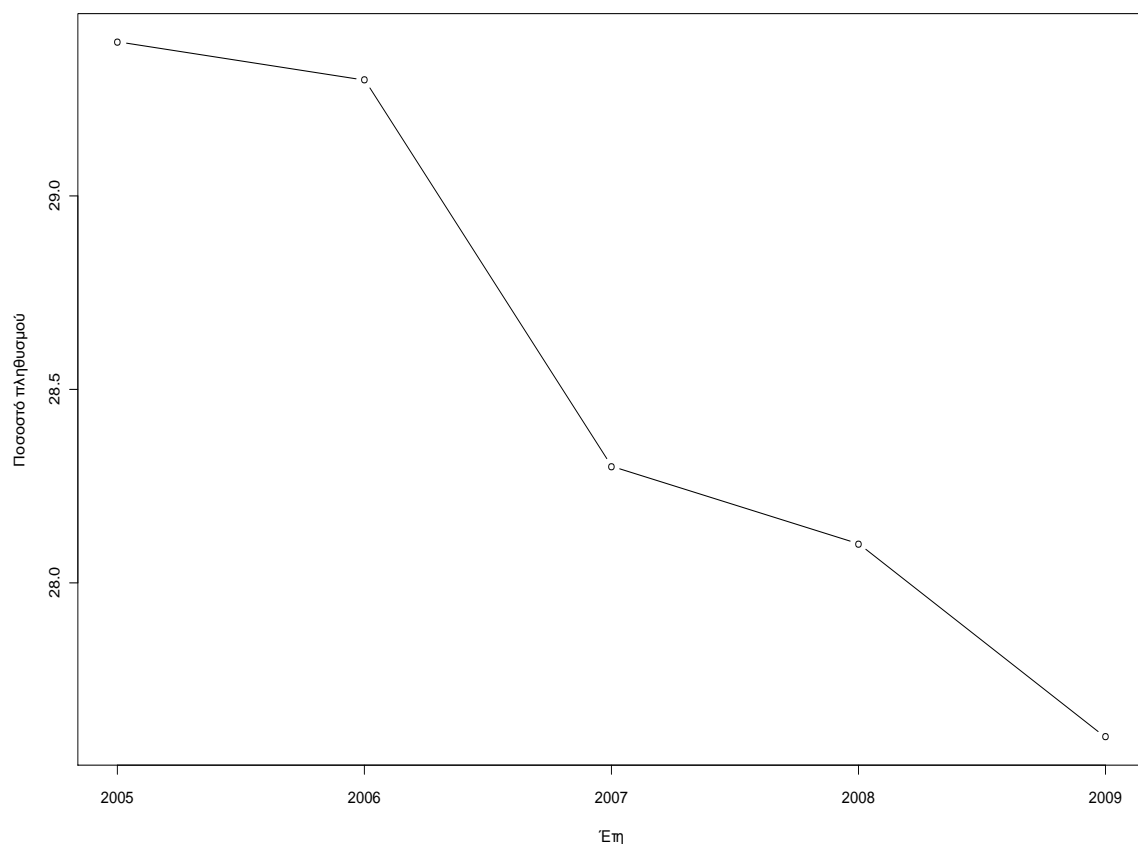
Ένωσης. Όσον αφορά το παράγοντα της ύπαρξης συγκρισιμότητας μπορούμε να δούμε ότι υπάρχει συγκρισιμότητα με την πάροδο του χρόνου αφού έχουμε ότι:

$$\Pi_{n-1} \leq \Pi_n + 0.3\% \text{ για όλα τα έτη.}$$

Η διαφορά της τάξεως του 0,3% που προέκυψε για τη τιμή του ποσοστού της χρονιάς n αποδεικνύει την ύπαρξη συγκρισιμότητας από έτος σε έτος και άρα τη μη ύπαρξη διαφορετικών στατιστικών διαδικασιών ή μεθόδων από έτος σε έτος.

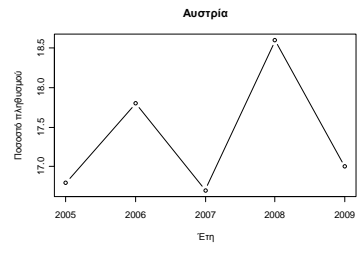
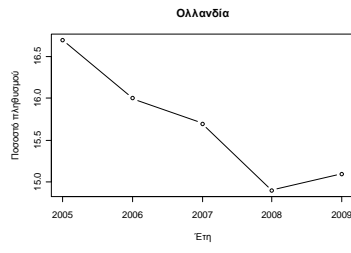
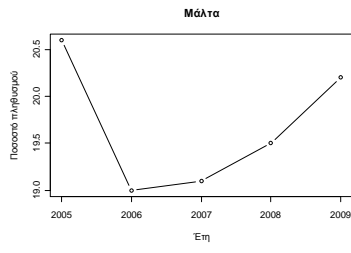
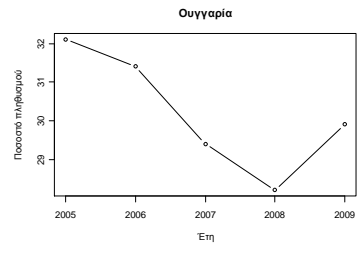
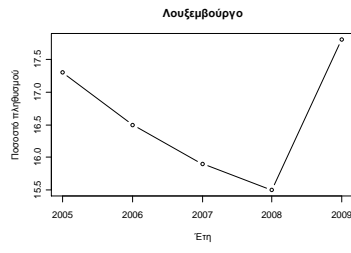
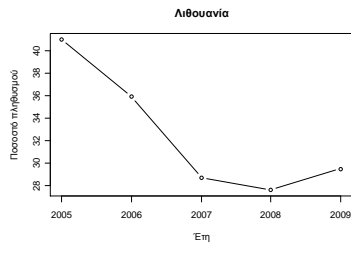
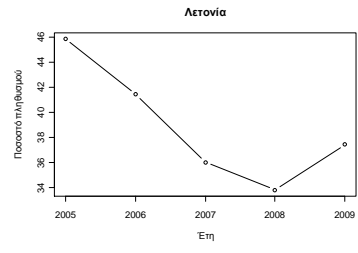
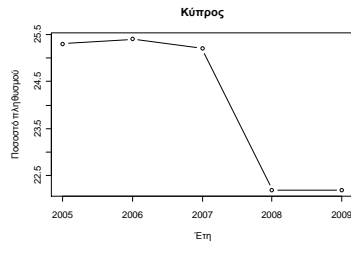
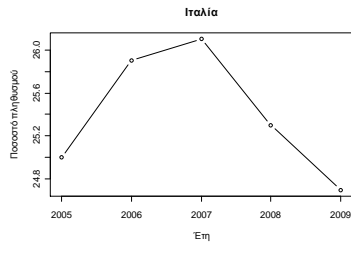
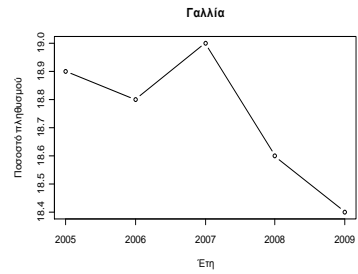
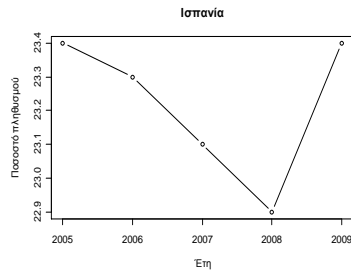
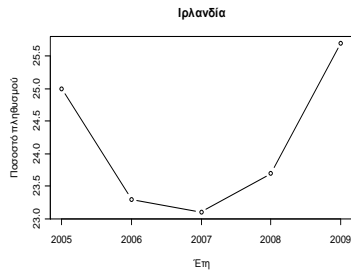
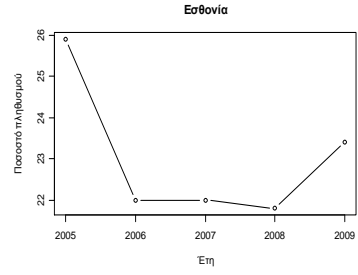
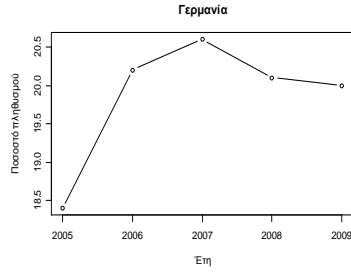
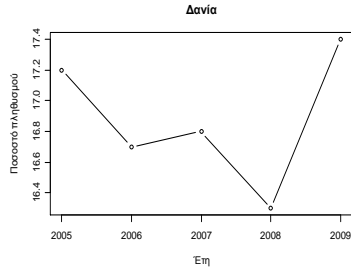
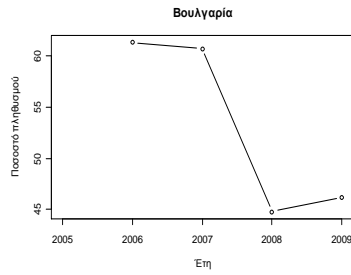
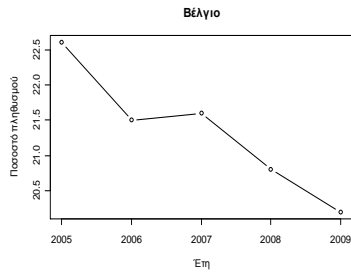
Στη συνέχεια, επαληθεύθηκε η ύπαρξη συγκρισιμότητας από όλες τις χώρες που συμμετέχουν στην έρευνα και για τις οποίες έχουν δημοσιευθεί αποτελέσματα για το ποσοστό του πληθυσμού τους στη φτώχεια για τα έτη 2005 έως 2009. Εν τούτοις, παρουσιάζεται αναλυτικά η περίπτωση της Ελλάδος και στη συνέχεια δίνονται οι γραφικές παραστάσεις όλων των χωρών.

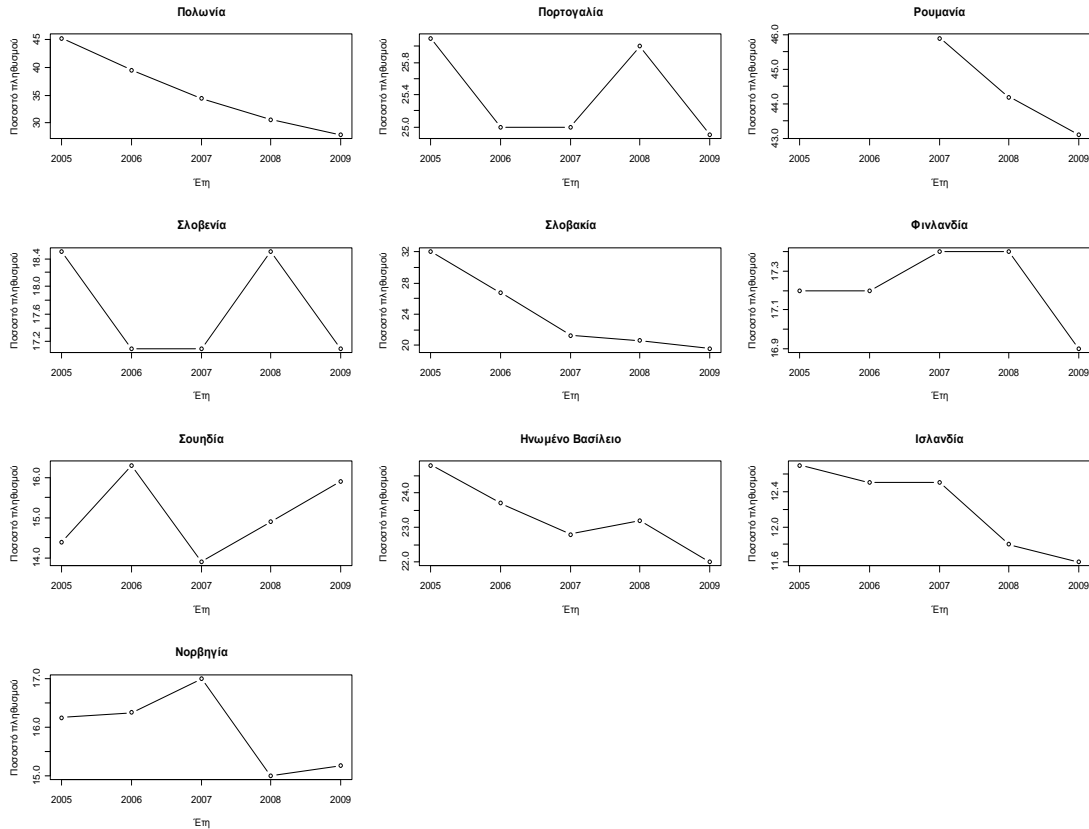
Το Ποσοστό πληθυσμού της Ελλάδος που παραμένουν στη φτώχεια.



Εύκολα προκύπτει ότι: $\Pi_{n-1} \leq \Pi_n + 0.5\%$ για όλα τα έτη, γεγονός που αποδεικνύει την ύπαρξη συγκρισιμότητας των ποσοστών με την πάροδο του χρόνου και για τη περίπτωση της Ελλάδος. Η τάξη μεγέθους θεωρείται, όπως και στις παραπάνω περιγραφές μικρή ώστε να οδηγεί στην ύπαρξη μη συγκρισιμότητας ανά έτος για τη μεταβλητή του ποσοστού της φτώχειας.

Παρακάτω δίνονται οι γραφικές παραστάσεις για τις υπόλοιπες χώρες για τις οποίες δημοσιεύονται αποτελέσματα ποσοστών πληθυσμού στη φτώχεια τη χρονική περίοδο 2005-2009.





Οι τιμές του ποσοστού φτώχειας τόσο για τη χώρα της Ελβετίας όσο και για τη χώρα της Τουρκίας δίνονται μόνο για το έτος 2009. Από τις γραφικές παραστάσεις με όμοια διαδικασία φαίνεται η δυνατότητα συγκρισιμότητας της μεταβλητής του ποσοστού φτώχειας.

Γεωγραφική Συγκρισιμότητα

Με σκοπό τη διασφάλιση συγκρισιμότητας τόσο των δεδομένων όσο και των δεικτών που μελετώνται η EU-SILC έχει υιοθετήσει μία στρατηγική εναρμονισμού των αποτελεσμάτων. Όταν χρησιμοποιείται στρατηγική εναρμονισμού των αποτελεσμάτων οι χώρες που λαμβάνουν μέρος στην έρευνα μπορούν να καθορίσουν το σχεδιασμό της και τις μεθόδους που θα χρησιμοποιήσουν έτσι ώστε να καλύπτονται οι απαιτήσεις για την επίτευξη εναρμονισμού των αποτελεσμάτων. Επιπλέον, οι χώρες καθορίζουν τον εθνικό σχεδιασμό της έρευνας και τις διαδικασίες επεξεργασίας των δεδομένων έτσι ώστε μέσω αυτών να απεικονίζεται ο διεθνής σχεδιασμός της έρευνας. Υπάρχουν στρατηγικές εναρμονισμού των αποτελεσμάτων, ο εκ των προτέρων εναρμονισμός και ο εκ των υστέρων. Στην περίπτωση του εκ των προτέρων εναρμονισμού οι χώρες μπορούν να υιοθετήσουν έρευνες που έχουν ήδη γίνει με σκοπό την παράδοση συγκρίσιμων αποτελεσμάτων ενώ στην περίπτωση του εκ των υστέρων εναρμονισμού οι έρευνες κατασκευάζονται από τις χώρες σύμφωνα με τα αποτελέσματα που απαιτείται να παραδοθούν.

Η EU-SILC, όπως εκτενώς έχει αναφερθεί στο Κεφάλαιο 1, βασίζεται σε ένα κοινό πλαίσιο καθορισμένο από εναρμονισμένες λίστες των βασικών και δευτερευόντων μεταβλητών, ένα κοινό προτεινόμενο σχεδιασμό της έρευνας, κοινές απαιτήσεις και ταξινόμηση με σκοπό τη βέλτιστη συγκρισιμότητα των πληροφοριών που συγκεντρώνονται. Το κοινό πλαίσιο όπως προειπώθηκε έχει αναφερθεί στο Κεφάλαιο 1. Βέβαια, αναφέρεται ότι το κοινό πλαίσιο δεν είναι αυστηρό όσον αφορά τις ιδιαιτερότητες της κάθε χώρας. Οι χώρες που συμμετέχουν στην έρευνα έχουν τη δυνατότητα να επαναπροσδιορίσουν κάποιο από τα χαρακτηριστικά του κοινού πλαισίου με σκοπό τη βελτίωση της ποιότητας των δεδομένων.

Παράρτημα:

Οι εντολές που απαιτήθηκαν στην γλώσσα R(συγκρισιμότητα- εισόδημα).

```
> data<-read.table("Income2.txt")
> data<-as.data.frame(data)
> names(data)<-c('Y2005','Y2006','Y2007','Y2008','Y2009')
> data1<-read.table("Income3.txt")
> data1<-as.data.frame(data1)
> x<-c(2005,2006,2007,2008,2009)
> names(data1)<-
c('EU27','EU15','Belgium','Bulgaria','CzechRepublic','Denmark','Germany','Estonia',
'Ireland','Greece','Spain','France','Italy','Cyprus','Latvia','Lithuania','Luxembourg',
'Hungary','Malta','Netherlands','Austria','Poland','Portugal','Romania','Slovenia',
'Slovakia','Finland','Sweden','UnitedKingdom','Iceland','Norway','Switzerland',
'Turkey')

> plot(x,data1$EU27,type="b",ylab="EU-27",xlab="Έτη",col="blue")
> title(main="Το ετήσιο εισόδημα των 27 κρατών μελών της Ευρωπαϊκής Ένωσης το
διάστημα 2005-2009.")
> data1$EU27
[1] 12975.2 13187.6 13900.4 14598.0 14622.0
> mean(data1$EU27)
[1] 13856.64
> var(data1$EU27)
[1] 590464.1

> plot(x,data1$EU15,type="b",ylab="EU-15",col="blue")
> title(main="Το ετήσιο εισόδημα των 15 κρατών μελών της Ευρωπαϊκής Ένωσης το
διάστημα 2005-2009.")
> data1$EU15
[1] 15465.2 15574.6 16545.2 17282.0 17290.0
> mean(data1$EU15)
[1] 16431.4
> var(data1$EU15)
[1] 785328.4

> plot(x,data1$Greece,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Το ετήσιο εισόδημα της Ελλάδος το διάστημα 2005-2009.")
> data1$Greece
[1] 9417.4 9850.0 10200.0 10800.0 11496.0
```



```

> par(mfrow=c(3,3))
> plot(x,data1$Belgium,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Βέλγιο")
> plot(x,data1$Bulgaria,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Βουλγαρία")
> plot(x,data1$CzechRepublic,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Τσεχική Δημοκρατία")
> plot(x,data1$Denmark,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Δανία")
> plot(x,data1$Germany,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Γερμανία")
> plot(x,data1$Estonia,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Εσθονία")
> plot(x,data1$Ireland,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Ιρλανδία")
> plot(x,data1$Spain,type="b",ylab="Ισπανία",xlab="Έτη")
> title(main="Ισπανία")
> plot(x,data1$France,type="b",ylab="Γαλλία",xlab="Έτη")
> title(main="Γαλλία")
> par(mfrow=c(3,3))
> plot(x,data1$Italy,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Ιταλία")
> plot(x,data1$Cyprus,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Κύπρος")
> plot(x,data1$Latvia,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Λετονία")
> plot(x,data1$Lithuania,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Λιθουανία")
> plot(x,data1$Luxembourg,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Λουξεμβούργο")
> plot(x,data1$Hungary,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Ουγγαρία")
> plot(x,data1$Malta,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Μάλτα")
> plot(x,data1$Netherlands,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Ολλανδία")
> plot(x,data1$Austria,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")
> title(main="Αυστρία")
> par(mfrow=c(4,3))
> plot(x,data1$Poland,type="b",ylab="Ετήσιο Εισόδημα",xlab="Έτη")

```

```

> title(main="Πολωνία")
> plot(x,data1$Portugal,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Πορτογαλία")
> plot(x,data1$Romania,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Ρουμανία")
> plot(x,data1$Slovenia,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Σλοβενία")
> plot(x,data1$Slovakia,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Σλοβακία")
> plot(x,data1$Finland,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Φινλανδία")
> plot(x,data1$Sweden,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Σουηδία")
> plot(x,data1$UnitedKingdom,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Ηνωμένο Βασίλειο")
> plot(x,data1$Iceland,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Ισλανδία")
> plot(x,data1$Norway,type="b",ylab="Ετήσιο Εισόδημα",xlab="Ετη")
> title(main="Νορβηγία")

```

Οι εντολές που απαιτήθηκαν στην γλώσσα R(συγκρισιμότητα- ποσοστό φτώχειας).

```

> datap<-read.table("Poverty.txt")
> datap<-as.data.frame(datap)
> names(datap)<-c('Y2005','Y2006','Y2007','Y2008','Y2009')
> data2<-read.table("Poverty2.txt")
> data2<-as.data.frame(data2)
> x<-c(2005,2006,2007,2008,2009)
> names(data2)<-
c('EU27','EU15','Belgium','Bulgaria','CzechRepublic','Denmark','Germany','Estonia','Ireland',
'Greece','Spain','France','Italy','Cyprus','Latvia','Lithuania','Luxembourg','Hungary',
'Malta','Netherlands','Austria','Poland','Portugal','Romania','Slovenia','Slovakia',
'Finland','Sweden','UnitedKingdom','Iceland','Norway','Switzerland','Turkey')
> plot(x,data2$EU27,type="b",ylab="EU-27",xlab="Ετη",col="blue")
> title(main="Το Ποσοστό πληθυσμού των 27 κρατών μελών της Ευρωπαϊκής Ένωσης που παραμένουν στη φτώχεια.")
> data2$EU27
[1] 26.0 25.0 24.5 23.6 23.1
> mean(data2$EU27)
[1] 24.44
> var(data2$EU27)

```

```
[1] 1.313
```

```
> plot(x,data2$EU15,type="b",ylab="EU-15",col="blue")  
> title(main="Το Ποσοστό πληθυσμού των 15 κρατών μελών της Ευρωπαϊκής  
Ένωσης που παραμένουν στη φτώχεια.")
```

```
> data2$EU15
```

```
[1] 21.5 21.8 21.7 21.5 21.1
```

```
> mean(data2$EU15)
```

```
[1] 21.52
```

```
> var(data2$EU15)
```

```
[1] 0.072
```

```
> plot(x,data2$Greece,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")  
> title(main="Το Ποσοστό πληθυσμού της Ελλάδος που παραμένουν στη φτώχεια.")
```

```
> data2$Greece
```

```
[1] 29.4 29.3 28.3 28.1 27.6
```

```
> mean(data2$Greece)
```

```
[1] 28.54
```

```
> var(data2$Greece)
```

```
[1] 0.613
```

```
> par(mfrow=c(3,3))
```

```
> plot(x,data2$Belgium,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Βέλγιο")
```

```
> plot(x,data2$Bulgaria,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Βουλγαρία")
```

```
> plot(x,data2$CzechRepublic,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Τσεχική Δημοκρατία")
```

```
> plot(x,data2$Denmark,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Δανία")
```

```
> plot(x,data2$Germany,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Γερμανία")
```

```
> plot(x,data2$Estonia,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Εσθονία")
```

```
> plot(x,data2$Ireland,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Ιρλανδία")
```

```
> plot(x,data2$Spain,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Ισπανία")
```

```
> plot(x,data2$France,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Γαλλία")
```

```
> par(mfrow=c(3,3))
```

```
> plot(x,data2$Italy,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
```

```
> title(main="Ιταλία")
```

```

> plot(x,data2$Cyprus,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Κύπρος")
> plot(x,data2$Latvia,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Λετονία")
> plot(x,data2$Lithuania,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Λιθουανία")
> plot(x,data2$Luxembourg,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Λουξεμβούργο")
> plot(x,data2$Hungary,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Ουγγαρία")
> plot(x,data2$Malta,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Μάλτα")
> plot(x,data2$Netherlands,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Ολλανδία")
> plot(x,data2$Austria,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Αυστρία")
> par(mfrow=c(4,3))
> plot(x,data2$Poland,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Πολωνία")
> plot(x,data2$Portugal,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Πορτογαλία")
> plot(x,data2$Romania,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Ρουμανία")
> plot(x,data2$Slovenia,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Σλοβενία")
> plot(x,data2$Slovakia,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Σλοβακία")
> plot(x,data2$Finland,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Φινλανδία")
> plot(x,data2$Sweden,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Σουηδία")
> plot(data2$UnitedKingdom,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Ηνωμένο Βασίλειο")
> plot(data2$Iceland,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Ισλανδία")
> plot(data2$Norway,type="b",ylab="Ποσοστό πληθυσμού",xlab="Έτη")
> title(main="Νορβηγία")

```

27 Κράτη - Μέλη της Ευρωπαϊκής Ένωσης	15 Κράτη - Μέλη της Ευρωπαϊκής Ένωσης
Αυστρία	Αυστρία
Βέλγιο	Βέλγιο
Βουλγαρία	
Κύπρος	
Τσεχική Δημοκρατία	
Δανία	Δανία
Εσθονία	
Φινλανδία	Φινλανδία
Γαλλία	Γαλλία
Γερμανία	Γερμανία
Ελλάδα	Ελλάδα
Ουγγαρία	
Ιρλανδία	Ιρλανδία
Ιταλία	Ιταλία
Λεττονία	
Λιθουανία	
Λουξεμβούργο	Λουξεμβούργο
Μάλτα	
Ολλανδία	Ολλανδία
Πολωνία	
Πορτογαλία	Πορτογαλία
Ρουμανία	
Σλοβακία	
Σλοβενία	
Ισπανία	Ισπανία
Σουηδία	Σουηδία
Ηνωμένο Βασίλειο	Ηνωμένο Βασίλειο

Βιβλιογραφία:

A) Ελληνική

Κοκολάκης, Γ. και Φουσκάκης, Δ. (2009). *Στατιστική: Θεωρία και Εφαρμογές*. Εκδόσεις Συμεών. Αθήνα.

Φουσκάκης, Δ. (2010). *Σημειώσεις Μαθήματος: Ανάλυση Δεδομένων με Η/Υ*. ΕΜΠ. Αθήνα.

Κοκολάκης, Γ. και Σπηλιώτης, Ι. (2002). *Εισαγωγή στις Πιθανότητες*. Εκδόσεις Συμεών. Αθήνα.

B) Διεθνής

Verma, V. and Betti, G. (2006). EU Statistics on income and living conditions (EU-SILC): Choosing the survey structure and sample design. *Statistics in Transition*, 7, pp. 935—970.

Gerstberger, C. (2010). *European Economic Statistics*. Eurostat. Statistical Books. Belgium.

Atkinson, A. and Marlier, E. (2010). *Income and living conditions in Europe*. Eurostat Statistical Books. Belgium.

Ehling, M. and Korner, T. (2007). *Handbook on Data Quality Assessments Methods and Tools*. Eurostat. Wiesbaden.

Jones, N. and Lewis D. (2003). *Handbook on improving quality by analysis of process variables*. Eurostat. Wiesbaden.

Fisenco, A. and Kozirkova V. (2006). Imputation in EU-SILC survey. Central Statistical Bureau of Latvia. Latvia.

Eurostat (2006). *Comparative EU Statistics on Income and Living Conditions: Issues and Challenges*. Office for Official Publications of the European Communities. Helsinki.

Eurostat (2010). *Robustness of some EU-SILC based indicators at regional level*. Office for Official Publications of the European Communities. Luxembourg.

Eurostat (2010). *The distribution of employees' labour earnings in the EU: data, concepts and first results*. Office for Official Publications of the European Communities. Luxembourg.

Eurostat (2010). *Income poverty and material deprivation in European Countries*. Office for Official Publications of the European Communities. Luxembourg.

Lelkes, O. and Zolyomi, E. (2008). *Poverty across Europe: The latest Evidence Using the EU-SILC survey*. European Centre Policy brief. Eisenbach.

Verma, V. (2001). *EU-SILC Sampling guidelines. Statistical Services in the Field of Income, Poverty and Social Exclusion and the New Instrument on Income and Living Conditions replacing the ECHP*. Office for Official Publications of the European Communities. Luxembourg.

Whelan, C. and Maitre, B. (2009). *Measuring Material Deprivation with EU-SILC: Lessons from the Irish Survey*. Economic and Social Research Institute. Dublin.

Eurostat (2003). *Methodological Documents Standard Report*. Office for Official Publications of the European Communities. Luxembourg.

Eurostat (2005). *The continuity of indicators during the transition between EHCP and EU-SILC*. Office for Official Publications of the European Communities. Luxembourg.

Museux, J. (2006). Weighting and estimation for the EU-SILC rotational design. *Eurostat and Statistics Finland International Conference on "Comparative EU Statistics on Income and Living Conditions: Issues and Challenges"*. Helsinki.

Eurostat (2009). *ESS Handbook for Quality Reports*. Office for Official Publications of the European Communities. Luxembourg.

Eurostat (2009). *ESS Standard for Quality Reports*. Office for Official Publications of the European Communities. Luxembourg.

Γ) Διαδικτυακή Πληροφόρηση

1. <http://epp.eurostat.ec.europa.eu/portal/page/portal/eurostat/home/>
2. <http://en.wikipedia.org>
3. http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/ilc_esms.htm
4. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2003:165:0001:0009:EN:PDF>
5. <http://stats.oecd.org/glossary/>
6. <http://www.statsoft.com/textbook/basic-statistics/>
7. www.efta.int
8. www.stats.oecd.org/glossary/
9. <http://ec.europa.eu>
10. http://epp.eurostat.ec.europa.eu/cache/ITY_SDDS/en/ilc_esms.htm#relevance
11. <http://www.eea.europa.eu/help/eea-help-centre/faqs/what-is-the-eu-15>
12. http://europa.eu/about-eu/member-countries/index_en.htm
13. http://epp.eurostat.ec.europa.eu/portal/page/portal/statistics/search_database