ΠΑΝΤΕΙΟΝ ΠΑΝΕΠΙΣΤΗΜΙΟ
ΚΟΙΝΩΝΙΚΩΝ ΚΑΙ ΠΟΛΙΤΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΤΜΗΜΑ ΚΟΙΝΩΝΙΚΗΣ ΠΟΛΙΤΙΚΗΣ

ΠΡΟΓΡΑΜΜΑ ΠΡΟΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ

Η ΧΡΗΣΗ ΜΙΚΡΟΔΕΔΟΜΕΝΩΝ ΤΗΣ ΒΑΣΗΣ IPUMSI ΓΙΑ ΤΗ ΣΤΑΤΙΣΤΙΚΗ ΑΝΑΛΥΣΗ ΔΗΜΟΓΡΑΦΙΚΩΝ ΚΑΙ ΚΟΙΝΩΝΙΚΩΝ ΧΑΡΑΚΤΗΡΙΣΤΙΚΩΝ ΤΟΥ ΠΛΗΘΥΣΜΟΥ ΤΗΣ ΕΛΛΑΔΑΣ, 2001

ΠΤΥΧΙΑΚΗ ΕΡΓΑΣΙΑ

Ελένη Κοσμά

Τριμελής εξεταστική επιτροπή:
Α. Μιχαλοπούλου, Αναπληρώτρια Καθηγήτρια, Εποπτεύουσα
Κλάιβ Ρίτσαρντσον, Καθηγητής
Μ. Συμεωνάκη, Λέκτορας

Αθήνα, Ιούνιος 2009

Copyright © Ελένη Κοσμά, 2009.
Με επιφύλαξη παντός δικαιώματος. All rights reserved.
Απαγορεύεται η αντιγραφή, αποθήκευση και διανομή της παρούσας εργασίας, εξ ολοκλήρου ή τμήματος αυτής, για εμπορικό σκοπό. Επιτρέπεται η ανατύπωση, αποθήκευση και διανομή για σκοπό μη κερδοσκοπικό, εκπαιδευτικής ή ερευνητικής φύσης, υπό την προϋπόθεση να αναφέρεται η πηγή προέλευσης και να διατηρείται το παρόν μήνυμα. Ερωτήματα που αφορούν τη χρήση της εργασίας για κερδοσκοπικό σκοπό πρέπει να απευθύνονται προς τον συγγραφέα.
Οι απόψεις και τα συμπεράσματα που περιέχονται σε αυτό το έγγραφο εκφράζουν τον συγγραφέα και δεν πρέπει να ερμηνευθεί ότι αντιπροσωπεύουν τις επίσημες θέσεις του Παντείου Πανεπιστημίου Κοινωνικών και Πολιτικών Επιστημών.

ΠΕΡΙΕΧΟΜΕΝΑ

ΕΙΣΑΓΩΓΗ ……..... 4

ΚΕΦΑΛΑΙΟ 1 ΜΕΘΟΔΟΛΟΓΙΑ ……... 6

1.1 Η διαδικασία επιλογής μεταβλητών …… 6

1.2 Σύγκριση των δημογραφικών και κοινωνικών χαρακτηριστικών του πληθυσμού της Ελλάδας της βάσης IPUMSI και της Απογραφής Πληθυσμού 2001 …… 9

ΚΕΦΑΛΑΙΟ 2 ΑΝΑΛΥΣΗ ΜΙΚΡΟΔΕΔΟΜΕΝΩΝ ……. 18

ΣΥΜΠΕΡΑΣΜΑΤΑ …… 21

ΠΗΓΕΣ-ΒΙΒΛΙΟΓΡΑΦΙΑ ……... 22

ΠΑΡΑΡΤΗΜΑ ……. 23

ΕΙΣΑΓΩΓΗ

Οι απογραφές πληθυσμού αποτέλεσαν τις πρώτες πηγές δημογραφικών και κοινωνικών δεδομένων. Όμως, διέθεταν σχεδόν αποκλειστικά αθροιστικά δεδομένα τα οποία δεν επέτρεπαν πρωτογενείς αναλύσεις. Για το λόγο αυτό, κατά τις τελευταίες δεκαετίες, έχει παρατηρηθεί σε αρκετές χώρες η ανάγκη συγκρότησης βάσεων δειγματοληπτικών δεδομένων από τις απογραφές πληθυσμού οι οποίες παρέχουν δεδομένα σε επίπεδο ατόμου (περίπτωσης), δηλαδή μικροδεδομένα, τα οποία είναι διαθέσιμα ελεύθερα στην ακαδημαϊκή κοινότητα. Αρχικά τη δεκαετία του 1960, οι ΗΠΑ δημιούργησαν την πρώτη βάση δειγματοληπτικών μικροδεδομένων στη Μινεσότα (Public Use Microdata Series). Τις δεκαετίες του 1970 και 1980 συγκρότησαν τέτοιες βάσεις ο Καναδάς και η Αυστραλία, αντίστοιχα. Ο μεγαλύτερος κίνδυνος σε αυτού του είδους τις βάσεις είναι η αποκάλυψη προσωπικών δεδομένων. Από σεβασμό προς τα συνεργαζόμενα κράτη και για την αποφυγή ταύτισης και αναγνώρισης ατόμων αφαιρούνται τα προσωπικά δεδομένα από τα άτομα, τις οικογένειες και τα νοικοκυριά ώστε να διασφαλίζεται η ανωνυμία της παρεχόμενης πληροφορίας.[1] Από το 1991, στην Μεγάλη Βρετανία συλλέχθηκαν μικροδεδομένα (Samples of Anonymized Records[2]) τα οποία διατίθενται ελεύθερα στην ακαδημαϊκή κοινότητα από τη βάση μικροδεδομένων του Πανεπιστημίου του Manchester. Ακολούθησαν και άλλες χώρες όπως η Ισπανία, το Λουξεμβούργο, η Γαλλία, η Ιρλανδία και η Ιταλία. Σε κάθε χώρα υπάρχει συνεργασία με την αντίστοιχη Στατιστική Υπηρεσία. Κάθε χώρα θέτει τους δικούς της περιορισμούς για τη διάθεση των μικροδεδομένων. Για παράδειγμα στην Ιταλία και το Λουξεμβούργο η πρόσβαση επιτρέπεται μόνο στην κυβέρνηση και όχι στην ακαδημαϊκή κοινότητα. Στις περισσότερες χώρες όμως η πρόσβαση είναι ελεύθερη με ορισμένους διακανονισμούς

[1] Βλ. σχετικά: Dale, A. και Elliot, M., "Proposals for 2001 samples of anonymised records: an assessment of disclosure risk", *JRSS*, A, 164 (*3*), (2001), http://www.jstor.org/stable/2680565, 427-447 (πρόσβαση στις 15-2-2009) και Gomatam, S., Karr, A.F., Reiter, J.P. and Sanil, A.P, "Data dissemination and disclosure limitation in a world without microdata: a risk-utility framework for remote access analysis servers", *Statistical Science*, 20 (*2*), (2005) 163-177, http://www.jstor.org/stable/20061168, (πρόσβαση στις 15-2-2009).

[2] Βλ. Samples of Anonymised Records, http://www.ccsr.ac.uk/sars/, (πρόσβαση στις 27-4-2009).

γιατί τα μικροδεδομένα είναι σημαντικά για τους ερευνητές, τους καθηγητές πανεπιστημίου καθώς και τους φοιτητές.[3]

Στο πλαίσιο αυτό συγκροτήθηκε και η διεθνής βάση μικροδεδομένων Integrated Public Use Microdata Series – International (IPUMSI) στο Minnesota Population Center, του Πανεπιστημίου της Minnesota.[4] Η ηλεκτρονική βάση δεδομένων IPUMSI είναι παγκοσμίως η μεγαλύτερη συλλογή με διαθέσιμα δεδομένα απογραφών πληθυσμού σε ατομικό επίπεδο. Τα δεδομένα είναι δείγματα από τις αντίστοιχες Απογραφές Πληθυσμού, σε σύνολο 44 χωρών που λαμβάνονται από το 1960 και μετά. Τα ονόματα και άλλες προσδιοριστικές πληροφορίες έχουν αφαιρεθεί. Στις μεταβλητές έχουν δοθεί οι κωδικοί και έχουν τεκμηριωθεί για να επιτρέψουν διεθνείς και διαχρονικές συγκρίσεις. Η βάση δεδομένων IPUMSI αποτελείται από μικροδεδομένα. Κάθε γραμμή του αρχείου αποτελεί ένα πρόσωπο με τα χαρακτηριστικά του. Το μεγαλύτερο πλήθος των προσώπων οργανώνονται σε νοικοκυριά, γεγονός που καθιστά δυνατή τη μελέτη των χαρακτηριστικών στο πλαίσιο των οικογενειών τους. Καθώς τα δεδομένα αναφέρονται σε άτομα, αυτό δίνει τη δυνατότητα στους ερευνητές να χρησιμοποιήσουν ένα στατιστικό πακέτο για την πρωτογενή ανάλυσή τους. Συγχρόνως μπορεί να γίνει η επιλογή των δειγμάτων και των μεταβλητών που απαιτούνται από κάθε χώρα για το έτος που ενδιαφέρει την ανάλυση. Η βάση μικροδεδομένων IPUMSI διαμορφώνεται σύμφωνα με το σύστημα δεδομένων IPUMS-USA, το οποίο περιέχει όλα τα δείγματα των απογραφών Πληθυσμού των Ηνωμένων Πολιτειών από το 1850 έως σήμερα, σε εθνικό επίπεδο. Η βάση δεδομένων διαμορφώθηκε στο Minnesota Population Center και συνεργάζεται με ένα πλήθος φορέων καθώς και οργανισμών από κάθε κράτος ώστε να συλλέξει δεδομένα. Στην Ελλάδα υπάρχει συνεργασία με την Εθνική Στατιστική Υπηρεσία Ελλάδος (ΕΣΥΕ). Συνεχώς γίνονται ανανεώσεις και επιτυγχάνεται προσθήκη νέων δεδομένων. Η πρόσβαση στη βάση δεδομένων IPUMSI δεν έχει χρηματική συνδρομή και είναι ελεύθερη στους ερευνητές και την ακαδημαϊκή κοινότητα.

[3] Βλ. σχετικά για τις πληροφορίες αυτές: Middleton, E., "Problems of harmonising UK-wide Samples of Anonymised Records from the 1991 Census", presented at IASSIS/IFDO 93 Conference held in Edinburgh, (1993), http://www.iassistdata.org/publicatio/iq/iq17/iqvol173-4middleton.pdf, 8, (πρόσβαση στις 27-4-2009).

[4] Βλ. Σχετικά: https://international.ipums.org/international/, (πρόσβαση στις 5-12-2008).

Η βάση IPUMSI είναι οργανωμένη σύμφωνα με τρεις βασικές αρχές: της εμπιστευτικότητας (confidentiality), της διαφύλαξης (preservation) και της εναρμόνισης (harmonization). Σύμφωνα με την αρχή της εμπιστευτικότητας τα συνεργαζόμενα έθνη τα οποία παραχωρούν τις απογραφές των πληθυσμών τους γνωρίζουν πως προστατεύεται η ταυτότητα των ατόμων, των οικογενειών και των νοικοκυριών. Επίσης, διαφυλάσσει τα μικροδεδομένα των απογραφών ιδιαίτερα εκείνα που κινδυνεύουν να καταστραφούν ή να χαθούν. Στη συνέχεια εναρμονίζει τα δεδομένα και την τεκμηρίωσή τους για να διευκολύνει τις συγκρίσεις με την παρέλευση του χρόνου και ανάμεσα στα έθνη, χωρίς απώλεια πληροφοριών. Αναλυτικότερα κάθε χώρα κωδικοποιεί διαφορετικά μια μεταβλητή. Η βάση δεδομένων δημιουργεί μία νέα μεταβλητή με την οποία επιτρέπονται οι συγκρίσεις μεταξύ της ίδιας μεταβλητής σε διαφορετικές χώρες. Πολλές μη εναρμονισμένες μεταβλητές χρησιμεύουν ως οι εισαγωγικές για τις εναρμονισμένες μεταβλητές IPUMSI. Αυτές μπορούν να χωριστούν σε δύο τύπους, τις γενικές και τις λεπτομερείς.

Το αντικείμενο της παρούσας εργασίας είναι η στατιστική ανάλυση των δημογραφικών και κοινωνικών χαρακτηριστικών του πληθυσμού της Ελλάδας για το 2001, σύμφωνα με τις μεταβλητές για το φύλο, την ηλικία, την οικογενειακή κατάσταση, την εργασία (απασχόληση) και την εκπαίδευση που παρέχονται από τη βάση μικροδεδομένων IPUMSI. Παράλληλα με την ανάλυση των μεταβλητών, και επειδή πρόκειται για δειγματοληπτικά δεδομένα, θα ελεγχθεί η ακρίβεια των κατανομών των μικροδεδομένων με τις αντίστοιχες δημοσιευμένες κατανομές από την Απογραφή Πληθυσμού 2001.[5] Στο πρώτο κεφάλαιο παρουσιάζεται η μεθοδολογία που ακολουθήθηκε, τα δεδομένα που χρησιμοποιήθηκαν από τη βάση μικροδεδομένων του IPUMSI και τη διαδικασία επιλογής τους. Αναφέρονται επίσης και τα αποτελέσματα των ελέγχων ακριβείας που διεξήχθησαν, δηλαδή τη σύγκριση των μεταβλητών που χρησιμοποιήθηκαν από τη βάση μικροδεδομένων καθώς και των αντίστοιχων μεταβλητών από την Απογραφή Πληθυσμού το 2001 για την Ελλάδα. Στο δεύτερο κεφάλαιο παρουσιάζεται η ανάλυση των δημογραφικών και κοινωνικών χαρακτηριστικών του πληθυσμού σύμφωνα με τα μικροδεδομένα για την Ελλάδα του 2001. Τέλος, υπό μορφή συμπερασμάτων παρουσιάζονται τα προβλήματα χρήσης δειγματοληπτικών δεδομένων κατά την ανάλυση.

[5] Βλ. ΕΣΥΕ, Απογραφή Πληθυσμού 2001, http://www.statistics.gr/StatMenu.asp.

ΚΕΦΑΛΑΙΟ 1 ΜΕΘΟΔΟΛΟΓΙΑ

1.1 Η διαδικασία επιλογής μεταβλητών

Τα μικροδεδομένα της βάσης IPUMSI διατίθενται από την ιστοσελίδα της βάσης στη διεύθυνση https://international.ipums.org. Η πρόσβαση στη βάση δεδομένων είναι απλή και η διαδικασία εγγραφής είναι αρκετά εύκολη. Σε πρώτο στάδιο θα χρειαστεί να γίνει μια εγγραφή (register), ώστε να καταγραφεί το ερευνητικό αίτημα στη βάση χρηστών που διαθέτει ο οργανισμός. Στην αίτηση αυτή (Apply for access) πρέπει να δηλωθεί η ιδιότητα του αιτούντος (καθηγητής, ακαδημαϊκός, ερευνητής, φοιτητής), ο λόγος για τον οποίο χρειάζονται τα δεδομένα, όπως επίσης και δήλωση άλλων δημογραφικών στοιχείων καθώς και στοιχείων επικοινωνίας. Μέσα από το ηλεκτρονικό ταχυδρομείο αποστέλλεται η επιβεβαίωση και γίνεται ενεργοποίηση του κωδικού πρόσβασης στη βάση. Σε δεύτερο στάδιο, στη διεύθυνση της βάσης IPUMSI γίνεται η εισαγωγή του κωδικού που έχει ήδη οριστεί, επιβεβαιώνεται η είσοδος στο σύστημα και η πρόσβαση στα δεδομένα. Στη συνέχεια, για τη δημιουργία ενός αρχείου δεδομένων, πρέπει να επιλεχθεί στην αρχική σελίδα η εντολή «create an extract» και αμέσως σας παραπέμπει στον κατάλογο με τις χώρες και τις απογραφές πληθυσμού, για τις οποίες υπάρχουν διαθέσιμα στοιχεία. Αφού κανείς επιλέξει τις χώρες και τις απογραφές πληθυσμού για τα έτη που τον/την ενδιαφέρουν (κάνοντας απλά επιλογή στα αντίστοιχα πεδία), στη συνέχεια πρέπει να προσδιοριστούν οι μεταβλητές που είναι απαραίτητες στην ανάλυση. Άμεσα δίνεται η εντολή και σε λίγες ώρες ή μέρες υπάρχουν διαθέσιμα στο ηλεκτρονικό ταχυδρομείο σε συμπιεσμένη μορφή. Τα αρχεία είναι διαθέσιμα για 72 ώρες. Σε αυτό το χρονικό διάστημα θα πρέπει να «κατέβουν» (download) από τον server σε συμπιεσμένη μορφή, η οποία μπορεί να γίνει με τη χρήση στατιστικών πακέτων (διατίθενται τα SPSS, STATA, SAS). Η διαδικασία για την αποσυμπίεση είναι λίγο περίπλοκη και εξαρτάται από το πακέτο που θα χρησιμοποιηθεί για την ανάλυση και έτσι θα πρέπει να ακολουθηθούν οι οδηγίες που βρίσκονται στο site.

Στην παρούσα εργασία, για την επεξεργασία των δεδομένων χρησιμοποιήθηκε το πακέτο στατιστικών προγραμμάτων SPSS. Οι μεταβλητές που χρησιμοποιήθηκαν είναι οι εξής: «SEX», «AGE», «MARST», «EDUCGR», και «EMPSTAT». Αυτές ορίζονται ως εξής[6]:

[6] Βλ. σχετικά τις μεταβλητές που παρουσιάζονται αναλυτικά στη σελίδα : https://international.ipums.org/international/resources/enum_materials_pdf/enum_form_gr2001.pdf.

- SEX: μέτρηση με βάση το φύλο του ατόμου.
- AGE: μέτρηση της ηλικίας με βάση το έτος γεννήσεως.
- MARST (marital status): η μέτρηση της οικογενειακής κατάστασης γίνεται με βάση το αν ο απογραφόμενος είναι άγαμος, έγγαμος, χήρος (-α), διαζευγμένος (-η) ή σε διάσταση.
- EDUCGR (education Greece): μέτρηση του επιπέδου εκπαίδευσης ανάλογα με το ανώτερο επίπεδο σπουδών που τελείωσε ο απογραφόμενος (δε θα δοθεί απάντηση για τα παιδιά που γεννήθηκαν μετά την 1/1/1995). Τα επίπεδα σπουδών είναι:

  1) Διδακτορικό
  2) Μάστερ
  3) Πτυχίο ΑΕΙ
  4) Πτυχίο ΤΕΙ, ΚΑΤΕ, ΚΑΤΕΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης
  5) Πτυχίο Μεταδευτεροβάθμιας Εκπαίδευσης (ΙΕΚ, Κολλέγια κλπ.)
  6) Απολυτήριο Γεν. Λυκείου ή 6τάξιου Γυμνασίου ή ΕΠΛ
  7) Πτυχίο ΤΕΛ
  8) Πτυχίο ΤΕΣ
  9) Απολυτήριο 3τάξιου Γυμνασίου
  10) Απολυτήριο Δημοτικού
  11) Φοιτά στο Δημοτικό
  12) Εγκατέλειψε το Δημοτικό, αλλά γνωρίζει γραφή και ανάγνωση
  13) Δε γνωρίζει γραφή και ανάγνωση
- EMPSTAT (employment status): η μεταβλητή καταγράφει την επαγγελματική κατάσταση του ατόμου. Μετριέται με βάση την κύρια ασχολία των απογραφόμενων την προηγούμενη εβδομάδα. Οι επιλογές είναι:

  1) Εργαζόμενος
  2) Ζητούσε εργασία
  3) Ζητούσε εργασία για πρώτη φορά
  4) Μαθητής (-τρια) ή σπουδαστής (-τρια)
  5) Συνταξιούχος
  6) Εισοδηματίας

7) Οικιακά

8) Στρατευμένος

9) Άλλη περίπτωση

Η βάση παρέχει εξαιρετική τεκμηρίωση για κάθε απογραφή και στην ιστοσελίδα της διατίθενται τα ερωτηματολόγια, οι ορισμοί εννοιών και μετρήσεων και οι κωδικοί των ταξινομήσεων, διευκολύνοντας την ανάλυση και τις διαχρονικές συγκρίσεις.

1.2 Σύγκριση των δημογραφικών και κοινωνικών χαρακτηριστικών του πληθυσμού της Ελλάδας της βάσης IPUMSI και της Απογραφής Πληθυσμού 2001

Τα δεδομένα της βάσης IPUMSI, όπως προαναφέραμε, είναι δειγματοληπτικά, δηλαδή βασίζονται σε ένα δείγμα νοικοκυριών από την Απογραφή Πληθυσμού του αντίστοιχου έτους, και περιλαμβάνονται όλα τα άτομα από κάθε επιλεγμένο στο δείγμα νοικοκυριό. Κατά συνέπεια, η χρήση τους στην ανάλυση δεν είναι τόσο απλή ή προφανής γιατί πρέπει να είναι γνωστά και να διατίθενται τα κατάλληλα βάρη.[7] Όμως, επειδή τα βάρη αυτά δεν διατίθενται, ο μόνος τρόπος για να ελεγχθεί η ακρίβεια των μετρήσεων του δείγματος και να υποδειχτούν σχετικά τα προβλήματα είναι να συγκριθούν οι αντίστοιχες κατανομές με αυτές της απογραφής.[8] Για να θεμελιωθεί η σύγκριση οι τιμές των μεταβλητών όπως προσφέρονται από τη βάση IPUMSI ανά-κωδικοποιήθηκαν σύμφωνα με πως δημοσιεύονται οι αντίστοιχες μετρήσεις στην Απογραφή Πληθυσμού.

Ένα άλλο σημείο που θα πρέπει να παρατηρήσουμε στην εργασία είναι πως η βάση δεδομένων IPUMSI συγκροτείται από δεδομένα απογραφών για την Ελλάδα το 2001 πραγματικού πληθυσμού και όχι μόνιμου πληθυσμού, αν και το γεγονός αυτό κατά την τεκμηρίωση της βάσης δεν διατυπώνεται σαφώς αλλά υπονοείται. (Για το λόγο αυτό, στο παράρτημα παρουσιάζουμε ενδεικτικά τα αποτελέσματα για το μόνιμο πληθυσμό.) Ως πληθυσμός στη βάση ορίζεται, όπως ακριβώς περιγράφεται στην ηλεκτρονική σελίδα IPUMSI, το σύνολο των ατόμων που βρέθηκαν στο

[7] Βλ. σχετικά: Michalopoulou, C.E., "Comparability Issues in the Greek Census Microdata, 1971-2001", paper presented at the IPUMS Workshop Athens, March 18, 20, 2008, http://www.hist.umn.edu/~rmccaa/ipums-europe/.

[8] Βλ. σχετικά: Tsimbos, C., Verropoulou, G. and Bagavos, C., "Native and Migrant Fertility Patterns in Greece: An Approach Based on IPUMS Microdata and Aggregated Census Data (1950-1969 cohorts)", paper presented at the IPUMS Workshop Athens, March 18, 20, 2008, http://www.hist.umn.edu/~rmccaa/ipums-europe/.

συγκεκριμένο τόπο την ημέρα της απογραφής και δεν συμπεριλαμβάνονται τα άτομα που είναι καταχωρημένα στα δημοτολόγια του αντίστοιχου δήμου. Έτσι, χαρακτηριστικά για τον πληθυσμού του 2001 μπορούμε να δούμε ότι περιλαμβάνει ολόκληρο τον πληθυσμό της χώρας, με σταθερό τόπο διαμονής καθώς και με μη καθορισμένο τόπο διαμονής.

Στον Πίνακα 1 παρουσιάζεται η σύγκριση των κατανομών του φύλου από την Απογραφή Πληθυσμού 2001 και των αντίστοιχων δεδομένων της βάσης IPUMSI. Παρατηρούμε πως οι αποκλίσεις στις κατανομές του φύλου δεν είναι σημαντικές και ανέρχονται σε 0,3 ποσοστιαίες μονάδες.

ΠΙΝΑΚΑΣ 1 Η σύγκριση των κατανομών του φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Φύλο	**Απογραφή Πληθυσμού* (%)**	**IPUMSI** (%)**	**Απόκλιση (%)**
Άντρες	49,5	49,2	0,3
Γυναίκες	50,5	50,8	0,3
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**

*N=10964020.
**N=1028884.

Στον Πίνακα 2 παρουσιάζεται η σύγκριση των κατανομών της ηλικίας από την Απογραφή του Πληθυσμού 2001 και των αντίστοιχων δεδομένων της βάσης IPUMSI.

ΠΙΝΑΚΑΣ 2 Η σύγκριση των κατανομών της ηλικίας από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Ηλικιακή ομάδα	**Απογραφή Πληθυσμού * (%)**	**IPUMSI** (%)**	**Απόκλιση (%)**
0-14	15,2	15,2	-
15-24	13,6	14,3	0,7
25-39	22,4	22,9	0,4
40-54	20,3	20,0	0,3
55-64	11,3	11,0	0,3
65-79	14,1	13,7	0,4
80+	3,0	3,0	-
ΣΥΝΟΛΟ	**100,0**	**100,1**	**-**

*N=10964020.
**N=1028884.

Από τα δεδομένα που παρουσιάζονται στον Πίνακα 2 παρατηρούμε ότι οι αποκλίσεις στις κατανομές της ηλικίας δεν είναι σημαντικές και κυμαίνονται κυρίως

μεταξύ 0,3 έως 0,4 ποσοστιαίες μονάδες. Η απόκλιση 0,7 ποσοστιαίων μονάδων που σημειώνεται για την κατηγορία από 15 έως 24 ετών, υποδεικνύει μία ελαφρά υπεροχή αυτής της κατηγορίας στα μικροδεδομένα, αλλά χωρίς τη δυνατότητα υπολογισμού δειγματοληπτικών σφαλμάτων, δεν γνωρίζουμε αν αυτό οφείλεται κατά τύχη ή στη δειγματοληψία, καθώς δεν μπορεί να εκτιμηθεί η επίδραση αυτής της απόκλισης στις εκτιμήσεις. Προσπαθώντας εμπειρικά να διαπιστώσουμε τις επιδράσεις στα αποτελέσματα μιας τέτοιας απόκλισης προχωρήσαμε σε μια διερεύνηση συγκρίνοντας τις κατανομές της ηλικίας σύμφωνα με το φύλο από την Απογραφή του Πληθυσμού 2001 και των αντίστοιχων δεδομένων της βάσης IPUMSI. Οι συγκρίσεις αυτές παρουσιάζονται στον Πίνακα 3. Από τη σύγκριση διαπιστώνεται ότι, ενώ οι αποκλίσεις στις κατανομές της ηλικίας για τις γυναίκες δεν είναι σημαντικές και κυμαίνονται από 0,3 έως 0,4 ποσοστιαίες μονάδες, το ίδιο δεν ισχύει για τους άντρες. Η μεγαλύτερη απόκλιση 1,1 ποσοστιαίων μονάδων διαπιστώνεται για την κατηγορία από 15 έως 24 ετών και μια απόκλιση 0,7 ποσοστιαίων μονάδων για την κατηγορία από 65 έως 69 ετών. Κατά συνέπεια, αν και θα μπορούσαμε με ασφάλεια να αποδώσουμε την απόκλιση του 0,7 που διαπιστώσαμε στην απλή κατανομή της ηλικίας (Πίνακας 2) αποκλειστικά και μόνο στους άντρες αυτής της κατηγορίας (Πίνακας 3), δεν είναι τόσο απλό όσο φαίνεται, υποδεικνύοντας ότι δεν τόσο προφανής η χρήση δειγματοληπτικών δεδομένων.

ΠΙΝΑΚΑΣ 3 Η σύγκριση των κατανομών της ηλικίας κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Ηλικιακή ομάδα	ΑΝΤΡΕΣ		Απόκλιση (%)	ΓΥΝΑΙΚΕΣ		Απόκλιση (%)
	Απογραφή Πληθυσμού * (%)	IPUMSI** (%)		Απογραφή Πληθυσμού* (%)	IPUMSI** (%)	
0-14	15,8	15,9	0,2	14,5	14,5	-
15-24	15,1	14,0	1,1	13,5	13,2	0,3
25-39	23,5	23,0	0,5	22,3	21,9	0,4
40-54	20,0	20,4	0,4	20,0	20,3	0,3
55-64	10,5	10,9	0,4	11,4	11,7	0,3
65-79	12,6	13,3	0,7	14,8	15,0	0,2
80+	2,5	2,6	0,1	3,5	3,5	-
ΣΥΝΟΛΟ	**100,0**	**100,1**		**100,0**	**100,1**	**-**

*$N_{ΑΑ}$=5427682, $N_{ΓΑ}$=5536338, όπου $N_{ΑΑ}$ και $N_{ΓΑ}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα δεδομένα της Απογραφής Πληθυσμού.

**$N_{ΑΙ}$=506129, $N_{ΓΙ}$=522755, όπου $N_{ΑΙ}$ και $N_{ΓΙ}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI (p=0,000).

Στο πλαίσιο αυτών των ελέγχων, στον Πίνακα 4 παρουσιάζεται η σύγκριση της κατανομής για την οικογενειακή κατάσταση από την Απογραφή Πληθυσμού το 2001 με τα αντίστοιχα μικροδεδομένα της βάσης IPUMSI.

ΠΙΝΑΚΑΣ 4 Η σύγκριση των κατανομών της οικογενειακής κατάστασης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Οικογενειακή κατάσταση	Απογραφή Πληθυσμού* (%)	IPUMSI** (%)	Απόκλιση (%)
Άγαμοι	39,6	38,9	0,7
Έγγαμοι	49,9	51,4	1,5
Διαζευγμένοι	3,3	2,7	0,6
Χήροι	7,0	7,0	-
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**

*N=10934097.
**N=1028884.

ΠΙΝΑΚΑΣ 5 Η σύγκριση των κατανομών της οικογενειακής κατάστασης κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Οικογενειακή κατάσταση	ΑΝΤΡΕΣ		Απόκλιση (%)	ΓΥΝΑΙΚΕΣ		Απόκλιση (%)
	Απογραφή Πληθυσμού* (%)	IPUMSI** (%)		Απογραφή Πληθυσμού* (%)	IPUMS** (%)	
Άγαμοι	44,6	43,6	1,0	34,8	34,3	0,5
Έγγαμοι	50,6	52,3	1,7	49,2	50,6	1,4
Διαζευγμένοι	2,4	1,9	0,5	4,3	3,5	0,8
Χήροι	2,3	2,2	0,1	11,7	11,6	0,1
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**	**100,0**	**100,0**	**-**

*$N_{ΑΑ}$=5413426, $N_{ΓΑ}$=5520671, όπου $N_{ΑΑ}$ και $N_{ΓΑ}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα δεδομένα της Απογραφής Πληθυσμού.
**$N_{ΑΙ}$=506129, $N_{ΓΙ}$=522755, όπου $N_{ΑΙ}$ και $N_{ΓΙ}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI (p=0,000).

Η σύγκριση των κατανομών της οικογενειακής κατάστασης (Πίνακας 4) ανέδειξε μια απόκλιση 1,5 ποσοστιαίων μονάδων για τους έγγαμους στα μικροδεδομένα της βάσης IPUMSI. Περαιτέρω διερεύνηση αυτής της απόκλισης σύμφωνα με το φύλο (Πίνακας 5) αναδεικνύει νέες αποκλίσεις όχι μόνο για την κατηγορία των εγγάμων αντρών και γυναικών (Πίνακας 5) για τους οποίους οι αποκλίσεις γίνονται 1,7 ποσοστιαίες μονάδες (έγγαμοι άντρες) και 1,4 ποσοστιαίες μονάδες (έγγαμες γυναίκες) αλλά, παράλληλα, αναδεικνύει και άλλες αποκλίσεις για τους άγαμους άντρες (1,5 ποσοστιαίες μονάδες) και τις διαζευγμένες γυναίκες (0,8 ποσοστιαίες μονάδες). Κατά συνέπεια, είναι πολύ δύσκολο εμπειρικά να αποδοθεί αυτή η

απόκλιση. Το γεγονός αυτό υποδεικνύει ότι, όταν διατίθενται δειγματοληπτικά δεδομένα χωρίς τα συνοδευτικά δειγματοληπτικά σφάλματα, δηλαδή τα κατάλληλα για κάθε μέτρηση βάρη, η προσπάθεια πιο λεπτομερούς διερεύνησης σε κατηγορίες μικρότερου μεγέθους περιπλέκουν την ανάλυση παρά διαφωτίζουν.

ΠΙΝΑΚΑΣ 6 Η σύγκριση των κατανομών του επιπέδου εκπαίδευσης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Επίπεδο εκπαίδευσης	**Απογραφή Πληθυσμού* (%)**	**IPUMS** (%)**	**Απόκλιση (%)**
Κάτοχοι διδακτορικού τίτλου	0,3	0,3	-
Κάτοχοι Μάστερ	0,5	0,5	-
Πτυχιούχοι Ανωτάτων Σχολών	8,3	7,8	0,5
Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης	3,1	4,5	1,4
Πτυχιούχοι μεταδευτεροβάθμιας εκπαίδευσης	3,5	3,3	0,2
Απόφοιτοι Μέσης εκπαίδευσης	22,3	20,7	1,6
Πτυχιούχοι ΤΕΛ Πτυχιούχοι ΤΕΣ	3,8	1,9	1,9
Απόφοιτοι 3τάξιου Γυμνασίου	11,3	10,6	0,7
Απόφοιτοι Δημοτικού	30,5	29,2	1,3
Φοιτούν στο Δημοτικό	6,4	12,1	5,7
Εγκατέλειψαν το Δημοτικό, αλλά γνωρίζουν γραφή και ανάγνωση	6,3	3,4	2,8
Δεν γνωρίζουν γραφή και ανάγνωση	3,6	5,9	2,3
ΣΥΝΟΛΟ	**100,0**	**100,2**	**-**

*N=10934097.
**N=1028884.

Στον Πίνακα 6 παρουσιάζεται η σύγκριση των κατανομών του επιπέδου εκπαίδευσης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI, αντίστοιχα. Οι τιμές της μεταβλητής προσδιορίζουν έντεκα (11) κατηγορίες. Οι κατηγορίες για το επίπεδο εκπαίδευσης που δίνονταν από τη βάση IPUMSI προσαρμόστηκαν στις κατηγορίες που είναι δημοσιευμένες από την Απογραφή Πληθυσμού, ώστε να μπορεί να θεμελιωθεί η σύγκριση μεταξύ τους.

Η σύγκριση, όπως προκύπτει από τον Πίνακα 6, ανέδειξε γενικά υψηλές αποκλίσεις. Οι αποκλίσεις αυτές είναι αναλυτικά της τάξεως των:

- 1,4 ποσοστιαίων μονάδων στην κατηγορία Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης,
- 1,6 ποσοστιαίων μονάδων στην κατηγορία Απόφοιτοι Μέσης εκπαίδευσης,
- 1,9 ποσοστιαίων μονάδων στην κατηγορία Πτυχιούχοι ΤΕΛ και ΤΕΣ,
- 0,7 ποσοστιαίων μονάδων στην κατηγορία Απόφοιτοι 3τάξιου Γυμνασίου,
- 1,3 ποσοστιαίων μονάδων στην κατηγορία Απόφοιτοι Δημοτικού,
- 5,7 ποσοστιαίων μονάδων στην κατηγορία φοιτούν στο Δημοτικό,
- 2,8 ποσοστιαίων μονάδων στην κατηγορία εγκατέλειψαν το Δημοτικό αλλά γνωρίζουν γραφή και ανάγνωση και
- 2,3 ποσοστιαίων μονάδων στην κατηγορία δεν γνωρίζουν γραφή και ανάγνωση.

Καθώς οι αποκλίσεις αυτές είναι υψηλές για τις αρκετές κατηγορίες και ιδιαίτερα για τις μικρότερου μεγέθους κατηγορίες που φοιτούν στο Δημοτικό ή το εγκατέλειψαν και των λειτουργικά αναλφάβητων προχωρήσαμε σε μια εμπειρική, πιο λεπτομερή διερεύνηση αυτών των υψηλών αποκλίσεων σύμφωνα με το φύλο που παρουσιάζεται στον Πίνακα 7.

ΠΙΝΑΚΑΣ 7 Η σύγκριση των κατανομών του επιπέδου εκπαίδευσης κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Επίπεδο εκπαίδευσης	ΑΝΤΡΕΣ		Απόκλιση (%)	ΓΥΝΑΙΚΕΣ		Απόκλιση (%)
	Απογραφή Πληθυσμού* (%)	IPUMS** (%)		Απογραφή Πληθυσμού* (%)	IPUMS** (%)	
Κάτοχοι διδακτορικού τίτλου	0,3	0,3	-	0,1	0,1	-
Κάτοχοι Μάστερ	0,6	0,6	-	0,4	0,4	-
Πτυχιούχοι Ανωτάτων Σχολών	8,9	8,3	0,6	7,8	7,2	0,6
Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπ/σης	3,5	6,0	2,5	2,8	3,6	0,8
Πτυχιούχοι μεταδευτεροβάθμιας εκπαίδευσης	3,1	2,9	0,2	3,9	3,7	0,2
Απόφοιτοι Μέσης εκπαίδευσης	21,4	19,6	1,8	23,2	21,7	1,5
Πτυχιούχοι ΤΕΛ Πτυχιούχοι ΤΕΣ	7,0	2,6	4,4	1,8	1,2	0,6
Απόφοιτοι 3τάξιου Γυμνασίου	12,5	11,6	0,9	10,0	9,4	0,6
Απόφοιτοι Δημοτικού	30,0	28,8	1,2	40,0	29,5	10,5
Φοιτούν στο Δημοτικό	6,6	11,0	4,4	6,0	13,0	7,0
Εγκατέλειψαν το Δημοτικό, αλλά γνωρίζουν γραφή και ανάγνωση	4,8	1,8	3,0	7,6	4,9	2,7
Δεν γνωρίζουν γραφή και ανάγνωση	1,9	6,0	3,0	5,2	5,6	0,4
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**	**100,0**	**100,0**	**-**

*N_{AA}= 5086359, $N_{ΓA}$= 5210439, όπου N_{AA} και $N_{ΓA}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα δεδομένα της Απογραφής Πληθυσμού.

**N_{AI}=506129, $N_{ΓI}$=522755, όπου N_{AI} και $N_{ΓI}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI (p=0,000).

Σύμφωνα με τα δεδομένα που παρουσιάζονται στον Πίνακα 7 παρατηρούμε ότι, στην κατηγορία Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης η απόκλιση για τους άνδρες είναι 2,5 ποσοστιαίες μονάδες έναντι 0,8

ποσοστιαίων μονάδων για τις γυναίκες. Στην κατηγορία Απόφοιτοι Μέσης εκπαίδευσης η απόκλιση για τους άνδρες είναι 1,8 ποσοστιαίες μονάδες έναντι 1,5 ποσοστιαίων μονάδων για τις γυναίκες. Υψηλή απόκλιση για τους άνδρες εμφανίζεται στην κατηγορία Πτυχιούχοι ΤΕΛ και ΤΕΣ με 4,4 ποσοστιαίες μονάδες σε αντίθεση με τις γυναίκες που η απόκλιση είναι 0,4 ποσοστιαίες μονάδες. Το ίδιο μπορούμε να παρατηρήσουμε και για τους Απόφοιτους 3τάξιου Γυμνασίου με την απόκλιση για τους άνδρες να είναι στις 0,9 ποσοστιαίες μονάδες και 0,6 για τις γυναίκες. Επίσης, στις κατηγορίες εγκατέλειψαν το Δημοτικό αλλά γνωρίζουν γραφή και ανάγνωση και δεν γνωρίζουν γραφή και ανάγνωση, η απόκλιση των ανδρών και στις δύο κατηγορίες είναι στις 3 ποσοστιαίες μονάδες και των γυναικών στις 2,7 και 0,4 ποσοστιαίες μονάδες, αντίστοιχα. Αντίθετα, για την απόκλιση 1,3 ποσοστιαίων μονάδων για τους αποφοίτους του Δημοτικού (Πίνακας 6) διαπιστώνεται ότι η απόκλιση για τους άνδρες είναι στις 1,2 ποσοστιαίες μονάδες και για τις γυναίκες στις 10,5 ποσοστιαίες μονάδες, εξαιρετικά υψηλή απόκλιση. Το ίδιο παρατηρούμε για την απόκλιση 5,7 ποσοστιαίων μονάδων για αυτούς που φοιτούν στο Δημοτικό (Πίνακας 6), για τους άνδρες η απόκλιση φτάνει τις 4,4 ποσοστιαίες μονάδες και για τις γυναίκες τις 7 ποσοστιαίες μονάδες. Σύμφωνα με τις αποκλίσεις που παρουσιάζονται στον Πίνακα 7, θα μπορούσαμε ίσως να αποδώσουμε κάποιες από τις μεγάλες αποκλίσεις στους άνδρες με εξαίρεση την εξαιρετικά εντυπωσιακή απόκλιση 10,5 ποσοστιαίων μονάδων για τις απόφοιτες Δημοτικού. Αυτό όμως δεν είναι, όπως προαναφέραμε τόσο απλό καθώς δεν είναι τόσο προφανής η χρήση δειγματοληπτικών δεδομένων.

Στον Πίνακα 8 παρουσιάζεται η σύγκριση των κατανομών της εργασιακής κατάστασης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI.

ΠΙΝΑΚΑΣ 8 Η σύγκριση των κατανομών της εργασιακής κατάστασης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Εργασιακή κατάσταση	**Απογραφή Πληθυσμού* (%)**	**IPUMS** (%)**	**Απόκλιση (%)**
Εργαζόμενοι	41,5	41,9	0,4
Άνεργοι	5,8	5,3	0,5
Μη ενεργοί	52,6	52,7	0,1
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**

*N=9887737.
**N=927066.

ΠΙΝΑΚΑΣ 9 Η σύγκριση των κατανομών της εργασιακής κατάστασης κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Εργασιακή κατάσταση	ΑΝΤΡΕΣ		Απόκλιση (%)	ΓΥΝΑΙΚΕΣ		Απόκλιση (%)
	Απογραφή Πληθυσμού* (%)	IPUMSI** (%)		Απογραφή Πληθυσμού* (%)	IPUMS** (%)	
Εργαζόμενοι	50,6	54,0	3,4	30,0	30,2	0,2
Άνεργοι	8,6	6,0	2,6	6,8	4,6	2,2
Μη ενεργοί	40,7	40,0	0,7	63,1	65,2	2,1
ΣΥΝΟΛΟ	**100,0**	**100,0**	-	**100,0**	**100,0**	-

*N_{AA}=4874982, $N_{ΓA}$=501275, όπου N_{AA} και $N_{ΓA}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα δεδομένα της Απογραφής Πληθυσμού.

**N_{AI}=506129, $N_{ΓI}$=522755, όπου N_{AI} και $N_{ΓI}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI (p=0,000).

Παρατηρούμε ότι, οι αποκλίσεις στις κατανομές της εργασιακής κατάστασης (Πίνακας 8) δεν είναι σημαντικές και κυμαίνονται από 0,1 έως 0,5 ποσοστιαίες μονάδες. Στον Πίνακα 9 παρουσιάζεται η σύγκριση των κατανομών της εργασιακής κατάστασης από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI ανά φύλο. Από τη σύγκριση διαπιστώνεται ότι, ενώ οι αποκλίσεις στις κατανομές της εργασιακής κατάστασης (Πίνακας 8) δεν είναι σημαντικές και κυμαίνονται από 0,1 έως 0,5 ποσοστιαίες μονάδες, το ίδιο δεν ισχύει στις κατανομές του Πίνακα 9. Στην κατηγορία των αντρών η μεγαλύτερη απόκλιση 3,4 ποσοστιαίων μονάδων διαπιστώνεται για την κατηγορία των εργαζομένων, μια απόκλιση 2,6 ποσοστιαίων μονάδων για την κατηγορία των ανέργων και 0,7 ποσοστιαίες μονάδες για την κατηγορία των μη ενεργών. Στις γυναίκες διαπιστώνεται μια απόκλιση 2,2 ποσοστιαίων μονάδων στην κατηγορία των ανέργων και μια απόκλιση 2,1 ποσοστιαίων μονάδων στην κατηγορία των μη ενεργών. Κατά συνέπεια, αν και θα μπορούσαμε με ασφάλεια να αποδώσουμε τις υψηλές αποκλίσεις περισσότερο στους άντρες, δεν γίνεται γιατί όπως προαναφέραμε δεν είναι τόσο προφανής η χρήση των δειγματοληπτικών δεδομένων.

ΚΕΦΑΛΑΙΟ 2 ΑΝΑΛΥΣΗ ΜΙΚΡΟΔΕΔΟΜΕΝΩΝ

Τα αποτελέσματα της μονομεταβλητής και διμεταβλητής ανάλυσης που παρουσιάστηκαν στο προηγούμενο κεφάλαιο στο πλαίσιο της σύγκρισης των μικροδεδομένων της βάσης IPUMSI και των αντίστοιχων από την Απογραφή Πληθυσμού 2001, προσφέρονται και για ανάλυση. Κατά συνέπεια, σύμφωνα με τις απλές κατανομές των μικροδεδομένων της βάσης IPUMSI για τα δημογραφικά και κοινωνικά χαρακτηριστικά για τον πληθυσμό της Ελλάδας του 2001 που παρουσιάζονται στο προηγούμενο κεφάλαιο (Πίνακες 1, 2, 4, 6 και 8) διαπιστώνεται ότι, οι γυναίκες υπερέχουν ελαφρώς έναντι των ανδρών (Πίνακας 1), είναι κυρίως ηλικίας από 25 έως 39 ετών (σύμφωνα με τη μέση, διάμεσο και επικρατούσα τιμή) και την ηλιακή ομάδα από 40 έως 54 ετών να κατέχει τη δεύτερη αριθμητικά θέση (Πίνακας 2), έγγαμοι (σύμφωνα με την επικρατούσα τιμή, βλ. Πίνακα 4), απόφοιτοι 3τάξιου Γυμνασίου (σύμφωνα με τη μέση και διάμεσο τιμή) ενώ η πολυπληθέστερη ομάδα είναι οι απόφοιτοι Δημοτικού (Πίνακας 6) και στην πλειονότητά τους μη οικονομικά ενεργοί (Πίνακας 8).

Σύμφωνα με τα δεδομένα της διμεταβλητής ανάλυσης των μικροδεδομένων της βάσης IPUMSI που παρουσιάζονται στο προηγούμενο κεφάλαιο (Πίνακες 3, 5, 7 και 9) διαπιστώνεται ότι, υπάρχει συνάφεια μεταξύ του φύλου και των μεταβλητών της ηλικίας, της οικογενειακής κατάστασης, της εκπαίδευσης και της απασχόλησης. Έτσι, όσο αυξάνεται η ηλικία οι γυναίκες είναι περισσότερες σε σχέση με τους άνδρες (Πίνακας 3). Η οικογενειακή κατάσταση των γυναικών διαφοροποιείται από αυτή των ανδρών, δηλαδή είναι περισσότερες γυναίκες παντρεμένες σε σχέση με τους άνδρες (Πίνακας 5). Επίσης, και το εκπαιδευτικό επίπεδο διαφοροποιείται μεταξύ ανδρών και γυναικών, δηλαδή είναι περισσότερες γυναίκες απόφοιτες Δημοτικού και μέσης εκπαίδευσης σε σχέση με τους άντρες (Πίνακας 7). Μεγάλη διαφοροποίηση παρατηρείται μεταξύ ανδρών και γυναικών σχετικά με την απασχόληση, με τη μεγάλη πλειονότητα των γυναικών να δηλώνει ως οικονομικά μη ενεργή σε αντίθεση από τους άνδρες με πάνω από τους μισούς να δηλώνουν ότι εργάζονται (Πίνακας 9).

Ας σημειωθεί ότι, τα αποτελέσματα της διμεταβλητής ανάλυσης που προαναφέραμε και τα οποία παρουσιάστηκαν στο προηγούμενο κεφάλαιο αναφέρονταν κυρίως στα ζητήματα συγκρίσεων μεταξύ των κατανομών των μικροδεδομένων από τη βάση IPUMSI και των αντίστοιχων δημοσιευμένων κατανομών από την Απογραφή Πληθυσμού. Όμως, οι δυνατότητες που προσφέρονται

από την ανάλυση των μικροδεδομένων είναι πολύ περισσότερες από αυτές της Απογραφής Πληθυσμού. Έτσι, μπορούμε να προχωρήσουμε σε περισσότερες στατιστικές επεξεργασίες από ό,τι προσφέρουν τα δημοσιευμένα δεδομένα.

Στον Πίνακα 10 παρουσιάζεται ο πίνακας συνάφειας για την απασχόληση και την ηλικία (*p*=0,000).

Πίνακας 10 Η απασχόληση ανά την ηλικία σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Ηλικιακή ομάδα	Εργαζόμενοι (%)	Άνεργοι (%)	Μη ενεργοί (%)
0-14	-	1,0	11,1
15-24	10,1	34,7	17,0
25-39	42,5	40,5	9,5
40-54	35,6	18,2	12,6
55-64	9,8	5,5	15,4
65-79	2,0	-	28,0
80+	-	-	6,4
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**

*N=927066, *p*=0,000.

Στον Πίνακα 10 μπορούμε να παρατηρήσουμε πως διαφοροποιείται η απασχόληση σύμφωνα με την ηλικία. Έτσι, διαπιστώνουμε ότι, οι εργαζόμενοι είναι κυρίως ηλικίας από 25 έως 39 ετών (42,5%) και 40 έως 54 ετών (35,6%), ηλικίες που συγκεντρώνουν τη συντριπτική πλειονότητα των εργαζομένων (78,1%) ενώ μόνο ένα 10,1% των νέων ηλικίας 15 έως 24 ετών εργάζεται. Αντίθετα η ανεργία πλήττει τους νέους, έτσι οι άνεργοι είναι κυρίως ηλικίας από 25 έως 39 ετών (40,5%) και 15 έως 24 ετών (34,7%), ηλικίες που συγκεντρώνουν τη συντριπτική πλειονότητα των ανέργων (75,2%), αλλά και η ηλικία από 40 έως 54 ετών πλήττεται από την ανεργία (18,2%). Οι μη οικονομικά ενεργοί είναι κυρίως ηλικίας από 65 έως 79 ετών.

Στον Πίνακα 11 παρουσιάζεται ο πίνακας συνάφειας για την οικογενειακή κατάσταση και την ηλικία (*p*=0,000).

Πίνακας 11 Η οικογενειακή κατάσταση ανά ηλικία σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Οικογενειακή κατάσταση	**0-14 (%)**	**15-24 (%)**	**25-39 (%)**	**40-54 (%)**	**55-64 (%)**	**65-79 (%)**	**80+ (%)**
Άγαμοι	100,0	92,5	36,8	8,0	4,5	4,2	4,0
Έγγαμοι	-	7,1	59,0	83,2	81,7	69,5	41,5
Διαζευγμένοι	-	0,3	3,5	5,7	3,6	2,1	1,4
Χήροι	-	0,1	0,6	3,1	10,2	24,2	53,0
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

*N=1028884, *p*=0,000.

Στον Πίνακα 11 διαπιστώνουμε πως διαφοροποιείται η οικογενειακή κατάσταση σύμφωνα με την ηλικία. Παρατηρούμε ότι, όπως αναμένεται, οι νέοι ηλικίας από 0 έως 14 ετών (100,0%) και από 15 έως 24 ετών (92,5%) είναι κυρίως άγαμοι, ενώ οι έγγαμοι είναι κυρίως ηλικίας από 25 έως 39 ετών (59,0%), από 40 έως 54 ετών (83,2%), από 55 έως 64 ετών (81,7%) και από 65 έως 79 ετών (69,5%). Οι ηλικίας από 80 ετών και άνω είναι κυρίως χήροι (53%) και έγγαμοι (41,5%).

Στον Πίνακα 12 παρουσιάζεται ο πίνακας συνάφειας για την εκπαίδευση και την ηλικία (*p*=0,000) και διαπιστώνουμε πως το επίπεδο εκπαίδευσης διαφοροποιείται με την ηλικία. Έτσι, όπως άλλωστε αναμένεται, η ηλικία από 0 έως 14 ετών φοιτούν κυρίως στο δημοτικό (40%), οι νέοι ηλικίας από 15 έως 24 ετών και από 25 έως 29 ετών είναι κυρίως απόφοιτοι μέσης εκπαίδευσης (37,4% και 32,6%, αντίστοιχα) ενώ οι μεγαλύτερες ηλικίες από 40 ετών και άνω είναι στην πλειονότητά τους απόφοιτοι δημοτικού.

Πίνακας 12 επίπεδο εκπαίδευσης ανά την ηλικία σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Επίπεδο εκπαίδευσης	**0-14 (%)**	**15-24 (%)**	**25-39 (%)**	**40-54 (%)**	**55-64 (%)**	**65-79 (%)**	**80+ (%)**
Κάτοχοι διδακτορικού τίτλου	-	-	0,3	0,5	0,3	0,2	0,1
Κάτοχοι Μάστερ	-	0,2	1,2	0,7	0,3	0,1	-
Πτυχιούχοι Ανωτάτων Σχολών	-	2,8	14,9	12,7	7,1	4,1	3,4
Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπ/σης	-	3,8	9,2	6,7	3,0	1,2	0,7
Πτυχιούχοι μεταδευτεροβάθμιας εκπαίδευσης	-	6,3	6,4	3,5	1,5	0,8	0,6
Απόφοιτοι Μέσης εκπαίδευσης	-	37,4	32,6	24,2	15,3	10,5	7,1
Πτυχιούχοι ΤΕΛ Πτυχιούχοι ΤΕΣ	-	5,2	4,0	1,3	0,4	0,1	0,1
Απόφοιτοι 3τάξιου Γυμνασίου	2,2	31,1	13,5	8,4	5,1	3,9	4,0
Απόφοιτοι Δημοτικού	19,3	11,4	15,9	38,7	51,5	44,4	38,3
Φοιτούν στο Δημοτικό	40,0	0,6	0,8	1,9	11,0	23,9	23,4
Εγκατέλειψαν το Δημοτικό, αλλά γνωρίζουν γραφή και ανάγνωση	-	1,1	1,1	1,4	4,4	10,8	22,3
Δεν γνωρίζουν γραφή και ανάγνωση	38,5	-	-	-	-	-	-
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**	**100,0**

*N=1028884, *p*=0,000.

Στον Πίνακα 13 παρουσιάζεται ο πίνακας συνάφειας για το εκπαιδευτικό επίπεδο και την οικογενειακή κατάσταση (*p*=0,000).

ΠΙΝΑΚΑΣ 13 Το επίπεδο εκπαίδευσης ανά την οικογενειακή κατάσταση σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Επίπεδο εκπαίδευσης	**Άγαμοι (%)**	**Έγγαμοι (%)**	**Διαζευγμένοι (%)**	**Χήροι (%)**
Κάτοχοι διδακτορικού τίτλου	0,1	0,3	0,6	0,1
Κάτοχοι Μάστερ	0,6	0,5	0,9	0,1
Πτυχιούχοι Ανωτάτων Σχολών	5,7	9,9	11,4	2,8
Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης	3,8	5,4	5,5	0,9
Πτυχιούχοι μεταδευτεροβάθμιας εκπαίδευσης	4,1	3,0	4,4	0,9
Απόφοιτοι Μέσης εκπαίδευσης	20,4	22,0	26,9	10,3
Πτυχιούχοι ΤΕΛ και Πτυχιούχοι ΤΕΣ	2,7	1,5	1,6	0,1
Απόφοιτοι 3τάξιου Γυμνασίου	14,1	8,8	10,5	4,0
Απόφοιτοι Δημοτικού	15,7	37,7	29,8	41,2
Φοιτούν στο Δημοτικό	16,5	7,6	5,2	23,2
Εγκατέλειψαν το Δημοτικό, αλλά γνωρίζουν γραφή και ανάγνωση	1,1	3,4	3,0	16,4
Δεν γνωρίζουν γραφή και ανάγνωση	15,1	-	-	-
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**	**100,0**

*N=1028884, *p*=0,000.

Σύμφωνα με τα δεδομένα που παρουσιάζονται στον Πίνακα 13 διαπιστώνεται ότι, το επίπεδο εκπαίδευσης διαφοροποιείται σύμφωνα με την οικογενειακή κατάσταση. Οι άγαμοι είναι κυρίως απόφοιτοι μέσης εκπαίδευσης, ενώ οι έγγαμοι είναι κυρίως απόφοιτοι δημοτικού και δευτερευόντως απόφοιτοι μέσης εκπαίδευσης. Η πλειονότητα των διαζευγμένων είναι απόφοιτοι δημοτικού και δευτερευόντως απόφοιτοι μέσης εκπαίδευσης, ενώ οι χήροι στην συντριπτική πλειονότητά τους είναι απόφοιτοι δημοτικού.

Στον Πίνακα 14 παρουσιάζεται ο πίνακας συνάφειας για την απασχόληση και την οικογενειακή κατάσταση (*p*=0,000).

ΠΙΝΑΚΑΣ 14 Η απασχόληση ανά την οικογενειακή κατάσταση σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Οικογενειακή κατάσταση	Εργαζόμενοι (%)	Άνεργοι (%)	Μη ενεργοί (%)
Άγαμοι	28,6	60,5	32,2
Έγγαμοι	65,9	34,1	52,4
Διαζευγμένοι	4,1	4,2	2,1
Χήροι	1,5	1,2	13,4
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**

*N=927066, *p*=0,000.

Στον Πίνακα 14 μπορούμε να παρατηρήσουμε πως διαφοροποιείται η απασχόληση σύμφωνα με την οικογενειακή κατάσταση. Διαπιστώνουμε ότι, οι εργαζόμενοι είναι στην συντριπτική πλειονότητά τους έγγαμοι, αντίθετα από τους ανέργους που στην συντριπτική πλειονότητά τους είναι άγαμοι. Οι οικονομικά μη ενεργοί είναι κυρίως έγγαμοι και δευτερευόντως άγαμοι.

Στον Πίνακα 15 παρουσιάζεται ο πίνακας συνάφειας για το επίπεδο εκπαίδευσης και την απασχόληση (*p*=0,000).

Στον Πίνακα 15 παρατηρούμε το πώς διαφοροποιείται το επίπεδο εκπαίδευσης με την απασχόληση. Οι εργαζόμενοι στην πλειονότητά τους είναι απόφοιτοι μέσης εκπαίδευσης και απόφοιτοι Δημοτικού με 26,6% και 25,6%, αντίστοιχα, και ένα 15,1% είναι πτυχιούχοι ανωτάτων σχολών. Οι άνεργοι είναι απόφοιτοι μέσης εκπαίδευσης (29,3%) και απόφοιτοι Δημοτικού (21,5%) και πτυχιούχοι ανωτάτων (98%). Οι οικονομικά μη ενεργοί είναι στην πλειονότητά τους απόφοιτοι Δημοτικού (38,8%), απόφοιτοι μέσης εκπαίδευσης (19,6%) και μόλις ένα 3,4% είναι πτυχιούχοι ανωτάτων σχολών.

ΠΙΝΑΚΑΣ 15 Το επίπεδο εκπαίδευσης ανά την απασχόληση σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI για το 2001*

Επίπεδο εκπαίδευσης	**Εργαζόμενοι (%)**	**Άνεργοι (%)**	**Μη ενεργοί (%)**
Κάτοχοι διδακτορικού τίτλου	0,6	0,2	0,1
Κάτοχοι Μάστερ	1,0	0,7	0,1
Πτυχιούχοι Ανωτάτων Σχολών	15,1	9,8	3,4
Πτυχιούχοι ΤΕΙ, ΚΑΤΕ, Ανώτερης Σχολής και Εκκλησιαστικής εκπαίδευσης	8,6	7,9	1,8
Πτυχιούχοι μεταδευτεροβάθμιας εκπαίδευσης	5,6	8,4	1,7
Απόφοιτοι Μέσης εκπαίδευσης	26,6	29,3	19,6
Πτυχιούχοι ΤΕΛ, Πτυχιούχοι ΤΕΣ	3,3	5,4	0,8
Απόφοιτοι 3τάξιου Γυμνασίου	10,6	13,6	12,4
Απόφοιτοι Δημοτικού	25,6	21,5	38,8
Φοιτούν στο Δημοτικό	2,0	1,8	15,1
Εγκατέλειψαν το Δημοτικό, αλλά γνωρίζουν γραφή και ανάγνωση	1,0	1,4	6,1
ΣΥΝΟΛΟ	**100,0**	**100,0**	**100,0**

*N=927066, *p*=0,000.

Οι παραπάνω στατιστικές επεξεργασίες εκτός από τα αποτελέσματα που προσφέρουν υποδεικνύουν κυρίως τις δυνατότητες αναλύσεων που επιδέχονται τα μικροδεδομένα. Βεβαίως, από τη φύση τους, τα δεδομένα αυτά υπάγονται στους περιορισμούς των μετρήσεών τους. Έτσι, με εξαίρεση τα δημογραφικά χαρακτηριστικά, τα δεδομένα για παράδειγμα για την απασχόληση από την Απογραφή Πληθυσμού δεν επαρκούν παρά μόνο για το έτος αναφοράς. Η απασχόληση είναι μια από τις πιο ευμετάβλητες κοινωνικές μεταβλητές και για το λόγο αυτό κανείς πρέπει να παρακολουθεί τα αποτελέσματα της Έρευνας Εργατικού Δυναμικού. Παρόλα αυτά, η ελεύθερη διάθεση αυτών των δεδομένων σε ατομικό επίπεδο προσφέρεται για διερευνήσεις και επεξεργασίες που η αθροιστική τους μορφή δεν επιτρέπει.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Οι συγκρίσεις των μεταβλητών μεταξύ των μικροδεδομένων της βάσης IPUMSI και της Απογραφής Πληθυσμού 2001 δεν αποκάλυψαν μεγάλες αποκλίσεις για τις υπό θεώρηση μεταβλητές, εκτός από το επίπεδο εκπαίδευσης. Όμως, επειδή ακριβώς πρόκειται για δειγματοληπτικά δεδομένα πρέπει να λαμβάνουμε υπόψη μας ότι δεν είναι εύκολο να αγνοηθούν αποκλίσεις ακόμα και των 0,7 ποσοστιαίων μονάδων οι οποίες, χωρίς τη δυνατότητα υπολογισμού των δειγματοληπτικών σφαλμάτων, δεν είναι εύκολο να αποδοθούν στην τύχη ή τη δειγματοληψία, και κατά συνέπεια δεν μπορεί να εκτιμηθεί η επίδρασή τους στα αποτελέσματα. Το γεγονός αυτό υποδείχτηκε στην προσπάθεια πιο αναλυτικής διερεύνησης των αποκλίσεων αυτών, που κατέδειξε ότι, όσο πιο λεπτομερώς προσπαθεί κανείς να ελέγξει τις διαφορές, τα δειγματοληπτικά δεδομένα αναδεικνύουν και άλλες που και αυτές με τη σειρά τους δεν μπορούν να ερμηνευθούν. Κατά συνέπεια, η χρήση μικροδεδομένων δεν είναι τόσο απλή ή προφανής, και υποδεικνύει την ανάγκη υπολογισμού των δειγματοληπτικών σφαλμάτων.

Εφόσον ληφθούν υπόψη τα ζητήματα που προκύπτουν κατά την επεξεργασία και ανάλυση δειγματοληπτικών δεδομένων, η ανάλυση των μικροδεδομένων της βάσης IPUMSI κατέδειξε τις δυνατότητες που προσφέρονται στην έρευνα για πρωτογενείς, λεπτομερείς αναλύσεις, τις οποίες δεν επιτρέπουν τα δεδομένα των Απογραφών Πληθυσμού.

ΠΗΓΕΣ - ΒΙΒΛΙΟΓΡΑΦΙΑ

ΠΗΓΕΣ

Ηλεκτρονικές διευθύνσεις

Integrated Public Use Microdata Series, University of Minnesota — International: https://international.ipums.org/international/ (πρόσβαση στις 5-12-2008).

Samples of Anonymised Records http://www.ccsr.ac.uk/sars/gettingstarted/ (πρόσβαση στις 27-4-2009)

ΕΣΥΕ, Απογραφή Πληθυσμού 2001, http://www.statistics.gr/StatMenu.asp.

ΒΙΒΛΙΟΓΡΑΦΙΑ

Michalopoulou, C.E., "Comparability Issues in the Greek Census Microdata, 1971-2001", paper presented at the IPUMS Workshop Athens, March 18, 20, (2008), http://www.hist.umn.edu/~rmccaa/ipums-europe/.

Tsimbos, C., Verropoulou, G. and Bagavos, C., "Native and Migrant Fertility Patterns in Greece: An Approach Based on IPUMS Microdata and Aggregated Census Data (1950-1969 cohorts)", paper presented at the IPUMS Workshop Athens, March 18, 20, (2008), http://www.hist.umn.edu/~rmccaa/ipums-europe/.

Middleton, E.,"Problems of harmonising UK-wide Samples of Anonymised Records from the 1991 Census", presented at IASSIS/IFDO 93 Conference held in Edinburgh, (1993), http://www.iassistdata.org/publicatio/iq/iq17/iqvol173-4middleton.pdf, (πρόσβαση στις 27-4-2009).

Dale, A. και Elliot, M., "Proposals for 2001 samples of anonymised records: an assessment of disclosure risk", *JRSS*, A, 164 (*3*), (2001) http://www.jstor.org/stable/2680565.

Gomatam, S., Karr, A.F., Reiter, J.P. and Sanil, A.P. (2005), "Data dissemination and disclosure limitation in a world without microdata: a risk-utility framework for remote access analysis servers", *Statistical Science*, 20 (*2*), http://www.jstor.org/stable/20061168.

ΠΑΡΑΡΤΗΜΑ

Στον Πίνακα Π.1 παρουσιάζεται η σύγκριση των κατανομών της ηλικίας του μόνιμου πληθυσμού από την Απογραφή Πληθυσμού 2001 και των αντίστοιχων δεδομένων της βάσης IPUMSI. Παρατηρούμε πως οι αποκλίσεις στις κατανομές της ηλικίας είναι ασήμαντες και κυμαίνονται από 0,3 έως 0,5 ποσοστιαίες μονάδες. Η απόκλιση 0,7 ποσοστιαίων μονάδων που σημειώνεται για την κατηγορία 15-24 ετών, υποδεικνύει μία ελαφριά υπεροχή αυτής της κατηγορίας στα μικροδεδομένα. Προσπαθώντας εμπειρικά να διαπιστώσουμε τις επιδράσεις στα αποτελέσματα μιας τέτοιας απόκλισης προχωρήσαμε σε μια διερεύνηση συγκρίνοντας τις κατανομές της ηλικίας ανά φύλο από την Απογραφή του Πληθυσμού 2001 και των αντίστοιχων δεδομένων της βάσης IPUMSI που παρουσιάζονται στον Πίνακα Π.2.

ΠΙΝΑΚΑΣ Π.1 Η σύγκριση των κατανομών της ηλικίας (μόνιμος πληθυσμός) κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Ηλικιακή ομάδα	**Απογραφή Πληθυσμού* (%)**	**IPUMSI** (%)**	**Απόκλιση (%)**
0-14	15,2	15,2	-
15-24	14,3	13,6	0,7
25-39	22,9	22,4	0,5
40-54	20,0	20,3	0,3
55-64	11,0	11,3	0,3
65-79	13,7	14,1	0,4
80+	3,0	3,0	-
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**

*N=10934097.
**N=1028884.

Παρατηρούμε πως οι αποκλίσεις στις κατανομές της ηλικίας για τις γυναίκες κυμαίνονται από 0,1 έως 0,4 ποσοστιαίες μονάδες. Στις κατανομές της ηλικίας των αντρών όμως διαπιστώνεται στην κατηγορία 15-24 ετών, η μεγαλύτερη απόκλιση που κυμαίνεται στις 1,1 ποσοστιαίες μονάδες στην κατηγορία 15-24 ετών και 0,7 στην κατηγορία 65-79 ετών. Θα μπορούσαμε να υποστηρίξουμε πως η απόκλιση των 1,1 ποσοστιαίων μονάδων (Πίνακας Π.2) οφείλεται αποκλειστικά στους άντρες αυτής της κατηγορίας (Πίνακας Π.1), κάτι που δεν είναι τόσο απλό όσο φαίνεται καθώς η χρήση των δειγματοληπτικών δεδομένων δεν είναι τόσο προφανής.

ΠΙΝΑΚΑΣ Π.2 Η σύγκριση των κατανομών της ηλικίας (μόνιμος πληθυσμός) κάθε φύλου από την Απογραφή Πληθυσμού 2001 και των μικροδεδομένων της βάσης IPUMSI

Ηλικιακή ομάδα	ΑΝΤΡΕΣ		Απόκλιση (%)	ΓΥΝΑΙΚΕΣ		Απόκλιση (%)
	Απογραφή Πληθυσμού* (%)	IPUMSI ** (%)		Απογραφή Πληθυσμού* (%)	IPUMSI ** (%)	
0-14	15,8	16,0	0,2	14,6	14,5	0,1
15-24	15,1	14,0	1,1	13,5	13,2	0,3
25-39	23,5	23,0	0,5	22,3	21,9	0,4
40-54	20,0	20,4	0,4	20,0	20,3	0,3
55-64	10,5	10,9	0,4	11,4	11,7	0,3
65-79	12,6	13,3	0,7	14,7	15,0	0,3
80+	2,5	2,6	0,1	3,6	3,5	0,1
ΣΥΝΟΛΟ	**100,0**	**100,0**	**-**	**100,0**	**100,0**	**-**

*N_{AA}=5413426. $N_{ΓA}$=5520671, όπου N_{AA} και $N_{ΓA}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα δεδομένα της Απογραφής Πληθυσμού.

**N_{AI}=506129. $N_{ΓI}$=522755, όπου N_{AI} και $N_{ΓI}$ αναφέρονται στο σύνολο των ανδρών και των γυναικών, αντίστοιχα, σύμφωνα με τα μικροδεδομένα της βάσης IPUMSI (p=0,000).