



**ΠΜΣ Εφαρμοσμένης Οικονομικής  
Τμήμα Οικονομικών Επιστημών  
Πανεπιστήμιο Θεσσαλίας**

**ΣΥΝΑΡΤΗΣΗ ΠΡΟΣΦΟΡΑΣ ΕΡΓΑΣΙΑΣ  
ΚΑΙ ΕΚΤΙΜΗΣΗ ΕΛΑΣΤΙΚΟΤΗΤΑΣ ΜΙΣΘΟΥ  
ΓΙΑ ΤΗΝ ΕΛΛΑΔΑ**

**ΑΝΑΣΤΑΣΙΟΣ Δ. ΜΟΣΧΟΣ**

**Επιβλέπων Καθηγητής: ΛΕΚΤΟΡΑΣ ΑΘΗΝΑ ΟΙΚΟΝΟΜΟΥ**

**ΒΟΛΟΣ 2012**

### **Υπεύθυνη δήλωση**

Βεβαιώνω ότι είμαι συγγραφέας αυτής της διπλωματικής εργασίας και ότι κάθε βοήθεια την οποία είχα για την προετοιμασία της, είναι πλήρως αναγνωρισμένη και αναφέρεται στη διπλωματική εργασία. Επίσης έχω αναφέρει τις όποιες πηγές από τις οποίες έκανα χρήση δεδομένων, ιδεών ή λέξεων, είτε αυτές αναφέρονται ακριβώς είτε παραφρασμένες. Επίσης βεβαιώνω ότι αυτή η πτυχιακή εργασία προετοιμάστηκε από εμένα προσωπικά ειδικά για τις απαιτήσεις του προγράμματος μεταπτυχιακών σπουδών στην Εφαρμοσμένη Οικονομική του Τμήματος Οικονομικών Επιστημών του Πανεπιστημίου Θεσσαλίας.

Βόλος, Μάιος 2012

## Ευχαριστίες

Θα ήθελα να ευχαριστήσω τον καθηγητή Ιωάννη Θεοδοσίου του Πανεπιστημίου του Aberdeen, UK, που μου παραχώρησε τη βάση δεδομένων SOCIOLD προκειμένου να πραγματοποιήσω την εμπειρική μου έρευνα. Επίσης θα ήθελα να ευχαριστήσω την κα. Οικονόμου για τη βοήθεια που μου πρόσφερε σε θέματα σχετικά με τη διπλωματική εργασία.

## ΠΕΡΙΕΧΟΜΕΝΑ

Περίληψη.....	4
Abstract.....	5
Κεφάλαιο 1: Εισαγωγή.....	6
Κεφάλαιο 2: Πού οφείλονται οι διαφορετικές εκτιμήσεις της προσφοράς εργασίας;.....	11
Κεφάλαιο 3: Εναλλακτικές προσεγγίσεις στην εκτίμηση της συνάρτησης προσφοράς εργασίας	
3.1 Η διαχρονική ( <i>lifetime</i> ) συνάρτηση προσφοράς εργασίας υπό το φάσμα του μοντέλου κύκλου-ζωής.....	19
3.2 Μεγιστοποίηση συνάρτησης χρησιμότητας υπό περιορισμούς και εμπειρική εκτίμηση.....	22
Κεφάλαιο 4. Το πρόβλημα της μεροληψίας από την επιλογή του δείγματος ( <i>sample selection bias</i> )	
4.1 Τι είναι η μεροληψία επιλογής δείγματος και που οφείλεται;.....	29
4.2 Η μεθοδολογία των δύο σταδίων του Heckman – μοντέλο Heckit.....	32
4.3 Πιθανά προβλήματα του μοντέλου Heckit.....	34
4.4 Διαγνωστικοί έλεγχοι υποθέσεων για την ύπαρξη της μεροληψίας επιλογής...	38
Κεφάλαιο 5. Εμπειρική εκτίμηση της συνάρτησης προσφοράς εργασίας	
5.1 Δεδομένα.....	43
5.2 Ανάλυση και περιγραφή των μεταβλητών	
5.2.1 Κύριες μεταβλητές.....	45
5.2.2 Δευτερεύουσες μεταβλητές.....	47
5.3 Εμπειρική ανάλυση	
5.3.1 Μοντέλο Heckman.....	50
5.3.2 Ελαστικότητα προσφοράς εργασίας.....	57
Κεφάλαιο 6. Συζήτηση-Σχολιασμός.....	64
Κεφάλαιο 7. Συμπεράσματα-Προτάσεις για περαιτέρω έρευνα.....	66
ΒΙΒΛΙΟΓΡΑΦΙΑ.....	68
ΠΑΡΑΡΤΗΜΑ Α.....	74
ΠΑΡΑΡΤΗΜΑ Β.....	76

## ΠΕΡΙΛΗΨΗ

Η μελέτη διερευνά τη προσφορά εργασίας στην Ελλάδα για το έτος 2004, για άνδρες και γυναίκες μεταξύ 45 και 65 ετών. Σε πρώτο στάδιο εκτιμάται μία εξίσωση μισθού με το μοντέλο Heckman και σε δεύτερο στάδιο εκτιμάται η εξίσωση προσφοράς εργασίας με το μοντέλο Tobit. Η εκτιμημένη ελαστικότητα βρέθηκε  $-0,26$  υποδηλώνοντας, πρώτον ότι το αποτέλεσμα εισοδήματος κυριαρχεί έναντι του αποτελέσματος υποκατάστασης και δεύτερον ότι η καμπύλη προσφοράς εργασίας είναι ανελαστική. Στη βιβλιογραφία, η ελαστικότητα προσφοράς εργασίας ανδρών και γυναικών διαφέρει σημαντικά, γι' αυτό οι περισσότερες έρευνες μελετούν μόνο τους άντρες ή μόνο τις γυναίκες. Επίσης σημαντικές διαφορές παρατηρούνται και μεταξύ νέων και μεγαλύτερων αντρών (Borjas, 2003). Οι διαφορές στις ελαστικότητες οφείλονται στην εξειδίκευση του μοντέλου, στη μέθοδο εκτίμησης (Evers et al., 2008) αλλά και στη χρονική περίοδο ορισμού των ωρών εργασίας (Klevmarken, 2005). Όσο μεγαλύτερη είναι η περίοδος ορισμού (πχ έτος), τόσο πιο ελαστική είναι η καμπύλη προσφοράς εργασίας (Borjas, 2003). Στη παρούσα μελέτη οι ώρες εργασίας είναι σε μηνιαία βάση, οπότε η ανελαστική εκτίμηση της προσφοράς εργασίας συμφωνεί με τη θεωρία.

## **ABSTRACT**

The study examines labor supply of Greek men and women, aged 45-65, for the year 2004. At first a wage equation is estimated using Heckman model and then a labor supply function is estimated using Tobit model. The estimated wage elasticity is found to be -0,26 suggesting that income effect dominates substitution effect and that labor supply curve is inelastic. Previous literature suggests that labor supply between men and women varies a lot, so they focus on either men or women. Many diversities are also found between younger and older cohorts for men (Borjas, 2003). The differences in the estimated elasticities are mainly due to model specification and estimation method which is applied (Evers et al., 2008), but also due to what time-period are the hours of work measured (Klevmarken, 2005). When the timespan is long (e.g. a year), the labor supply curve tends to be more elastic (Borjas, 2003). In the present study, hours of work are measured in monthly base so the rather inelastic labor supply estimation agrees with the theory.

# Κεφάλαιο 1

## Εισαγωγή

Η θεωρία της αγοράς εργασίας είναι ένα πολύ μεγάλο κεφάλαιο της οικονομικής επιστήμης και χωρίζεται σε δύο μέρη. Το ένα είναι η ζήτηση εργασίας από τη πλευρά των εργοδοτών και το άλλο η προσφορά εργασίας από τη πλευρά των εργαζομένων. Στη παρούσα μελέτη εξετάζεται η προσφορά εργασίας. Βασικό στοιχείο της θεωρίας προσφοράς εργασίας είναι τα δύο είδη μισθών με τα οποία έρχονται αντιμέτωποι οι εργαζόμενοι. Ο ένας είναι ο μισθός επιφύλαξης (*reservation wage*) και ο άλλος ο μισθός εργασίας (*labor wage*). Ο μισθός επιφύλαξης αντικατοπτρίζει εκείνο το επίπεδο μισθού στο οποίο ένα άτομο είναι πρόθυμο να εργαστεί, ενώ ο μισθός εργασίας αντιστοιχεί στο επίπεδο μισθού που απολαμβάνει ένα άτομο από την εργασία του. Ένα άτομο δηλαδή, θα συμμετέχει στην αγορά εργασίας μόνο όταν ο μισθός εργασίας είναι μεγαλύτερος ή τουλάχιστον ίσος με τον μισθό επιφύλαξης. Όταν ο μισθός εργασίας που του προσφέρεται είναι μικρότερος από τον μισθό επιφύλαξης, τότε θα προτιμήσει να μην εργαστεί (Borjas, 2003: σελ. 70-71; Birch, 2005; Maglad, 1998; El-Hamidi, 2003; Heckman, 1993; Heckman and MaCurdy, 1980).

Σε αντιστοιχία με τα δύο είδη μισθών, υπάρχουν και δύο πλευρές στις οποίες διακρίνεται η μελέτη της προσφοράς εργασίας. Οι όροι που χρησιμοποιούνται στη βιβλιογραφία για τον διαχωρισμό αυτόν είναι *extensive margin* και *intensive margin*. Το *extensive margin* αφορά την απόφαση που παίρνει κάποιος για το αν θα εισέλθει στην αγορά εργασίας (αν εργαστεί) ή αν μείνει εκτός. Δηλαδή συνδέεται άμεσα με τον μισθό επιφύλαξης και τον μισθό εργασίας. Αντίθετα, το *intensive margin* αφορά την απόφαση που παίρνει κάποιος για το πόσες ώρες θα εργαστεί, δεδομένου ότι έχει πάρει την απόφαση να εισέλθει στην αγορά εργασίας (Kuroda and Yamamoto, 2008; Shaw, 1992). Οι περισσότερες έρευνες μελετούν το *intensive margin*, μέσω της κλασικής συνάρτησης προσφοράς εργασίας με εξαρτημένη μεταβλητή τις ώρες εργασίας και ανεξάρτητη το μισθό. Κάποιες όμως μελετούν και το *extensive margin* που ως εξαρτημένη μεταβλητή έχει τον ρυθμό συμμετοχής (*participation rate*), και ανεξάρτητες το μισθό και διάφορους άλλους παράγοντες που επηρεάζουν τη πιθανότητα κάποιος να εισέλθει στην αγορά εργασίας. Οι περισσότερες μεταβολές της προσφοράς εργασίας οφείλονται στον αριθμό των ατόμων που εισέρχονται ή εξέρχονται από την αγορά εργασίας. Δηλαδή οι μεγαλύτερες αλλαγές στη συμπεριφορά των ατόμων παρατηρούνται στο *extensive margin*, και όχι στο *intensive* (Heckman, 1993; Kuroda and Yamamoto, 2008).

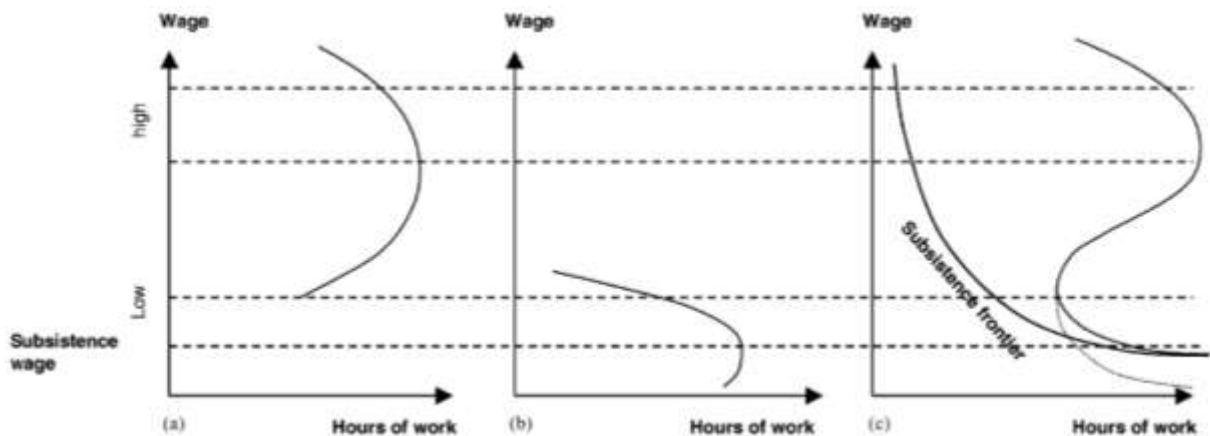
Οι ελαστικότητες προσφοράς εργασίας που συνήθως εκτιμώνται, είναι η ελαστικότητα ως προς τον μισθό (ελαστικότητα μισθού) και η ελαστικότητα ως προς το εισόδημα (εισοδηματική ελαστικότητα), καθώς και οι αντίστοιχες ‘σταυρωτές’ ελαστικότητες, δηλαδή ως προς τον μισθό και το εισόδημα του/της συζύγου. Τελευταίο σημείο στο οποίο πρέπει να γίνει αναφορά, αναφορικά με τους διαχωρισμούς στην εμπειρική μελέτη της προσφοράς εργασίας, είναι τα δύο είδη ελαστικοτήτων που εκτιμώνται. Οι όροι που χρησιμοποιούνται είναι compensated και uncompensated ελαστικότητες. Οι compensated ελαστικότητες έχουν να κάνουν με τη χικσιανή συνάρτηση ζήτησης, η οποία αναφέρεται μόνο στο αποτέλεσμα υποκατάστασης και αγνοεί το αποτέλεσμα εισοδήματος, ενώ οι uncompensated ελαστικότητες έχουν να κάνουν με τη μαρσαλιανή συνάρτηση ζήτησης η οποία περιλαμβάνει το αποτέλεσμα εισοδήματος. Αν γνωρίζουμε τη μαρσαλιανή (uncompensated) ελαστικότητα και την εισοδηματική ελαστικότητα, τότε μπορεί να υπολογιστεί η χικσιανή (compensated) ελαστικότητα μέσω της εξίσωσης του Slutsky. Η σχέση που ισχύει είναι: μαρσαλιανή ελαστικότητα – συνολική εισοδηματική ελαστικότητα = χικσιανή ελαστικότητα (Mizala et al., 1999; El-Hamidi, 2003). Η μελέτη των Aaberge et al. (1999) είναι από τις πολύ λίγες μελέτες που περιλαμβάνει την εκτίμηση όλων των παραπάνω είδη ελαστικοτήτων (μαρσαλιανές και χικσιανές, ατομικές και σταυρωτές, extensive και intensive margin).

Η νεοκλασική θεωρία υποστηρίζει ότι η συνάρτηση προσφοράς εργασίας έχει το σχήμα  $\cap$  (Εικόνα 1α). Δηλαδή μέχρι το σημείο όπου η καμπύλη έχει θετική κλίση, το αποτέλεσμα υποκατάστασης είναι ισχυρότερο από αυτό του εισοδήματος. Εν τέλει όμως επικρατεί το αποτέλεσμα εισοδήματος και η καμπύλη στρέφει προς τα πίσω (Borjas, 2003: σελ. 73). Σύμφωνα όμως με την Dessing (2002) η νεοκλασική θεωρία έχει κάποιες αδυναμίες. Η Dessing (2002) εστιάζεται στις φτωχές οικογένειες και στη στρατηγική συμπεριφορά των μελών της για διαφορετικά επίπεδα μισθών. Υποστηρίζει ότι στην πραγματικότητα η καμπύλη προσφοράς εργασίας είναι της μορφής S (Εικόνα 1c). Δεν θεωρεί ότι το νεοκλασικό μοντέλο είναι λάθος, απλά είναι ατελές, αφού δεν λαμβάνει υπόψη τη πραγματική συμπεριφορά των φτωχών ανθρώπων. Σύμφωνα με το νεοκλασικό μοντέλο οι φτωχοί δουλεύουν ελάχιστα, όμως στη πράξη παρατηρείται ότι στις λιγότερο ανεπτυγμένες χώρες δουλεύουν περισσότερο όσο μειώνονται οι αποδοχές τους, προκειμένου να διατηρήσουν το εισόδημά τους σταθερό. Αυτό συμβαίνει ακόμα και στις πλούσιες χώρες σε περιόδους ύφεσης και μειωμένης οικονομικής δραστηριότητας. Με λίγα λόγια στη νεοκλασική θεωρία οι ελαστικότητες προσφοράς εργασίας (σε χαμηλά εισοδήματα) έχουν συνήθως θετικό πρόσημο, ενώ στην πραγματικότητα και σύμφωνα με τη συγγραφέα είναι αρνητικές. Η



συγγραφέας στηρίζεται σε παλιότερες μελέτες που έχουν γίνει σε φτωχές αγροτικές περιοχές, που θεωρούν ότι οι πάμφτωχοι-εξαθλιωμένοι άνθρωποι δουλεύουν (αναγκαστικά) πολλές ώρες με σκοπό να αποκτήσουν ένα επίπεδο εισοδήματος που θα τους εξασφαλίσει τις βασικές ανάγκες διαβίωσης. Μόλις αποκτήσουν αυτό το εισόδημα διαβίωσης, ή εναλλακτικά (Berg, 1961) το εισόδημα-στόχος (*income target*), μειώνουν τις ώρες εργασίας σε κάθε επιπλέον αύξηση του μισθού και αυτό γιατί δεν τους ενδιαφέρει να συσσωρεύσουν περαιτέρω πλούτο. Τέτοια είδους συμπεριφορά χαρακτηρίζεται ως ‘τεμπέλικη-υψηλής σχόλης’ στη βιβλιογραφία και τα άτομα αυτά έχουν χαμηλές φιλοδοξίες, αφού στόχο έχουν να δουλέψουν τόσο ώστε να βγάλουν τα προς το ζην. Σχηματικά η προσφορά εργασίας αυτών των ανθρώπων παρουσιάζει μία ‘κάμψη προς τα πίσω’ (*backward-bending*).

Εικόνα 1: (a) canonical model (b) backward-bending at low wages (c) S-shaped labor supply curve



Πηγή: Dessing (2002)

Η Εικόνα 1 δείχνει συνοπτικά όλες αυτές τις εναλλακτικές απόψεις που έχουν επικρατήσει στη βιβλιογραφία όσον αφορά τη μορφή των καμπυλών προσφοράς εργασίας. Στο τμήμα (a) είναι το νεοκλασικό μοντέλο προσφοράς εργασίας, στο τμήμα (b) η ‘προς τα πίσω κάμψη’ σε χαμηλά επίπεδα μισθού, όπως περιγράφηκε προηγουμένως, και στο τμήμα (c) έχουμε τη καμπύλη προσφοράς σε σχήμα (αντίστροφου) S, η οποία μπορούμε να πούμε ότι είναι μία συγχώνευση του τμήματος (a) και (b). Αυτό που προκύπτει γενικά και από τα τρία διαγράμματα είναι ότι η ελαστικότητα προσφοράς εργασίας είναι αρνητική στα υψηλά και χαμηλά επίπεδα μισθού και θετική στα ενδιάμεσα τμήματα. Επίσης, από το πρώτο και το τρίτο διάγραμμα φαίνεται καθαρά αυτό που υποστηρίζει η συγγραφέας, ότι δηλαδή η νεοκλασική θεωρία αδυνατεί να αντιμετωπίσει τη συμπεριφορά των χαμηλόμισθων-φτωχών

ανθρώπων. Η κοίλη μορφή του νεοκλασικού μοντέλου υποδηλώνει ότι μέχρι το σημείο καμπής υπερσχύει το αποτέλεσμα υποκατάστασης (το αποτέλεσμα υποκατάστασης είναι θετικό και το αποτέλεσμα εισοδήματος αρνητικό), ενώ από το σημείο καμπής και όσο παραπέρα αυξάνεται το επίπεδο του μισθού, υπερσχύει το αποτέλεσμα εισοδήματος.

Η όλη ανάλυση και μελέτη της συγγραφέως στηρίζεται στο τρίτο διάγραμμα το οποίο περιλαμβάνει την S-καμπύλη προσφοράς εργασίας και το σύνορο διαβίωσης. Το σύνορο διαβίωσης αντικατοπτρίζει ουσιαστικά το αναγκαίο εισόδημα που πρέπει να αποκτήσει κάποιος προκειμένου να μπορέσει να συντηρήσει τον εαυτό του και την οικογένειά του<sup>1</sup>. Με λίγα λόγια κάτω από το σύνορο διαβίωσης, η οικογένεια δεν μπορεί να επιβιώσει. Η ανοιχτόχρωμη γραμμή στο τμήμα (c) είναι μία εναλλακτική μορφή της καμπύλης προσφοράς εργασίας, όπως περιγράφηκε από τον Sharif (1991) μελετώντας την Ινδία, με την υπόθεση όμως ότι δεν υπάρχει κάποιος περιορισμός διαβίωσης (*subsistence constraint*).

Συγκεκριμένα ο Sharif υποστήριζε ότι στα χαμηλά εισοδήματα η καμπύλη προσφοράς εργασίας 'πέφτει προς τα εμπρός και κάτω' (*forward falling*) και δεν είναι ακριβώς το ίδιο με τη 'κάμψη προς τα πίσω'. Η έρευνά του στηρίζεται στη μεγιστοποίηση μίας συνάρτησης χρησιμότητας που αποτελείται από δύο αγαθά, τροφή και ξεκούραση, και χαρακτηρίζεται ως συνάρτηση χρησιμότητας με αυξανόμενη ελαστικότητα υποκατάστασης μεταξύ των δύο αγαθών. Στη συνέχεια, μέσω των εκτιμημένων παραμέτρων της συνάρτησης χρησιμότητας, κατασκευάζει τη συνάρτηση προσφοράς και προχωράει στην εκτίμηση της καμπύλης προσφοράς εργασίας. Οι ελαστικότητες που βρίσκει είναι αρνητικές, όπως υποστηρίζει και η Dessing (2002).

Σκοπός της παρούσας μελέτης είναι να διερευνηθεί η συμπεριφορά των εργαζομένων στην Ελλάδα μέσω της συνάρτησης προσφοράς εργασίας και να εκτιμηθεί η ελαστικότητα προσφοράς εργασίας ως προς τον μισθό. Η εισοδηματική ελαστικότητα δεν μπορεί να εκτιμηθεί καθώς δεν υπάρχουν αρκετές παρατηρήσεις για τη μεταβλητή του πλούτου του νοικοκυριού.

Το υπόλοιπο μέρος της εργασίας έχει ως εξής: Στα κεφάλαια 2 και 3 γίνεται μία βιβλιογραφική ανασκόπηση και παρουσιάζονται τα θεωρητικά και εμπειρικά ευρήματα διάφορων μελετών. Στο κεφάλαιο 4 αναλύεται λεπτομερώς το μοντέλο του Heckman, τα προβλήματά του και κάποιοι έλεγχοι που γίνονται για τον εντοπισμό της μεροληψίας του δείγματος (*sample selection bias*), καθώς είναι το βασικό μοντέλο στο οποίο στηρίζεται η εμπειρική μας ανάλυση. Στο κεφάλαιο 5 παρουσιάζεται η εμπειρική ανάλυση, στο κεφάλαιο

---

<sup>1</sup> Το τμήμα (c) της Εικόνας 1 μπορεί να αντικατοπτρίζει τη καμπύλη προσφοράς εργασίας μόνο ενός ατόμου, αλλά και τη συνδυασμένη καμπύλη προσφοράς της οικογένειας.

6 κάποια συνοπτικά σχόλια και στο κεφάλαιο 7 τα συμπεράσματα και κάποιες καταληκτικές παρατηρήσεις.

## Κεφάλαιο 2

### Που οφείλονται οι διαφορετικές εκτιμήσεις των ελαστικοτήτων προσφοράς εργασίας;

Οι Evers et al. (2008) κάνοντας μία ανασκόπηση της υπάρχουσας βιβλιογραφίας που αφορά τις εμπειρικές εκτιμήσεις ελαστικοτήτων προσφοράς εργασίας, εξετάζουν τους παράγοντες που επηρεάζουν συστηματικά τα μεγέθη των ελαστικοτήτων. Ακόμη, διερευνούν τους λόγους στους οποίους αποδίδονται οι διαφορές και οι παρεκκλίσεις των μελετών όσον αφορά τα αριθμητικά μεγέθη των ελαστικοτήτων διαφορετικών χωρών. Έτσι συγκέντρωσαν συνολικά 30 μελέτες δημοσιευμένες σε ευρέως γνωστά ακαδημαϊκά περιοδικά από τις οποίες πήραν 209 εκτιμημένες ελαστικότητες. Στις 30 αυτές μελέτες δεν θέλησαν να συμπεριλάβουν όσες είχαν εκτιμήσει τις ελαστικότητες με τη μέθοδο των ελαχίστων τετραγώνων, καθώς σύμφωνα με τη βιβλιογραφία η μέθοδος OLS δίνει μη συνεπείς εκτιμήσεις.

Κύριος σκοπός τους ήταν να δουν κατά πόσο οι μέθοδοι εκτίμησης καθώς και η εξειδίκευση των μοντέλων (*model specification*) επηρεάζουν το μέγεθος των ελαστικοτήτων. Γι' αυτό το λόγο εκτός από τις συνήθεις μεταβλητές που είχαν χρησιμοποιηθεί στις 30 έρευνες και αφορούν βασικά χαρακτηριστικά που παίζουν ρόλο στην εκτίμηση των ελαστικοτήτων (όπως πχ ο μισθός, ώρες εργασίας, οικογενειακή κατάσταση, ηλικία, επίπεδο μόρφωσης κλπ), οι συγγραφείς κατασκεύασαν και δικές τους μεταβλητές-ψευδομεταβλητές που αφορούν τις χώρες προέλευσης των μελετών, τις μεθόδους εκτίμησης (πχ 2SLS, Maximum Likelihood, non-parametric) και την εξειδίκευση της συνάρτησης προσφοράς εργασίας (πχ γραμμική, τετραγωνική). Στη συνέχεια χώρισαν το δείγμα σε άντρες και γυναίκες και τα συμπεράσματα στα οποία κατέληξαν γενικά είναι, ότι και η οικονομετρική μέθοδος εκτίμησης που θα χρησιμοποιηθεί αλλά και η διαφορετική εξειδίκευση της συνάρτησης προσφοράς εργασίας φαίνεται να επηρεάζουν τις εκτιμήσεις των ελαστικοτήτων. Για παράδειγμα βρήκαν ότι στους άντρες μία ημι-λογαριθμική συνάρτηση δίνει υψηλότερες ελαστικότητες από οποιαδήποτε άλλη εξειδίκευση του υποδείγματος, ενώ στις γυναίκες συμβαίνει ακριβώς το αντίθετο. Επίσης η μέθοδος μεγίστης πιθανοφάνειας παράγει γενικά μικρές εκτιμήσεις. Η διαπίστωση αυτή έρχεται σε αντίθεση με τους MaCurdy et al. (1990), οι οποίοι ισχυρίστηκαν ότι η μέθοδος μεγίστης πιθανοφάνειας παράγει σχετικά υψηλές εκτιμήσεις ελαστικοτήτων. Όμως, σύμφωνα με τους MaCurdy et al. (1990), οι υψηλές

ελαστικότητες οφείλονται στην υιοθέτηση του μοντέλου Hausman<sup>2</sup>, το οποίο στηρίζεται σε ισχυρούς περιορισμούς<sup>3</sup>. Αν όμως δεν υιοθετείται το μοντέλο Hausman τότε οι ελαστικότητες θα είναι μικρότερες. Τέλος οι Evers et al. (2008) προχωρούν σε προβλέψεις για διάφορες χώρες (Αγγλία, ΗΠΑ, Σουηδία, Ολλανδία) και βρίσκουν ότι οι ελαστικότητες των γυναικών είναι μεγαλύτερες από των αντρών σε όλες τις υπό εξέταση χώρες, και συγκεκριμένα στην Ολλανδία είναι γύρω στο 0,1 για τους άντρες και 0,5 για τις γυναίκες.

Εκτός όμως από την εξειδίκευση του υποδείγματος και την μεθοδολογία εκτίμησης, σημαντικό ρόλο παίζει και ο τρόπος συλλογής των δεδομένων. Η άντληση των πληροφοριών ως γνωστόν, γίνεται επί των πλείστων με ερωτήσεις δια του τηλεφώνου ή δια προσώπου και με τη διανομή και συλλογή ερωτηματολογίων. Οι ώρες εργασίας, που είναι και η σημαντικότερη μεταβλητή στον υπολογισμό των ελαστικοτήτων, εύκολα μπορεί να μετρηθεί λανθασμένα (Klevmarken, 2005). Εξαρτάται από τον τρόπο διατύπωσης της ερώτησης από τον ερευνητή αλλά και από την προσωπική απάντηση του εργαζομένου. Συνήθως όταν ένας εργαζόμενος ερωτάται σχετικά με το πόσες ώρες δουλεύει τη μέρα, την εβδομάδα, τον μήνα κλπ, η απάντηση που δίνει είναι οι ώρες εργασίας που δηλώνει το συμβόλαιο/η σύμβασή του. Οι ώρες συμβολαίου όμως, πιθανόν να μην αντιστοιχούν και στις πραγματικές ώρες εργασίας. Πχ το συμβόλαιο μπορεί να λέει ότι θα δουλεύει από 8 ώρες τη μέρα, αλλά στη πραγματικότητα μπορεί να δουλεύει 7 ώρες και την υπόλοιπη 1 ώρα να τη περνάει σε διαλείμματα για ξεκούραση και μικρογεύματα. Επομένως στις ώρες εργασίας σαν μεταβλητή μπορεί να περιλαμβάνονται και οι ώρες διαλείμματος. Όταν ένας ερευνητής θέλει να διαχωρίσει τις πραγματικές ώρες εργασίας από τις συμβατικές, μπορεί να το κάνει με το να διατυπώνει τις σωστές ερωτήσεις (πχ Πόσες ώρες δουλεύετε την εβδομάδα; Πόσες ώρες από αυτές αφιερώνετε σε διαλείμματα; κ.ο.κ). Εκτός όμως από αυτό το πρόβλημα υπάρχουν και άλλα: Όταν τα δεδομένα καλύπτουν μεγάλη χρονική περίοδο (πχ ετήσια δεδομένα) τότε ο εργαζόμενος δεν μπορεί να θυμάται ακριβώς πόσες ώρες δούλεψε μέσα σε αυτό το διάστημα, ή αν έκανε υπερωρίες, ούτε μπορεί να θυμάται αν αρρώστησε, πόσες μέρες έμεινε εκτός δουλειάς μέχρι να αναρρώσει, έκτακτα περιστατικά που τυχόν προέκυψαν κάποιες μέρες με αποτέλεσμα να λείπει για λίγο από τη δουλειά κ.α. Είναι προφανές λοιπόν ότι όλοι αυτοί οι παράγοντες μπορούν να δημιουργήσουν σφάλματα μέτρησης της προσφοράς εργασίας.

Ο Klevmarken (2005) θέλησε να δείξει ότι ο διαφορετικός τρόπος μέτρησης των ωρών εργασίας επηρεάζει το μέγεθος των ελαστικοτήτων. Γι' αυτό το λόγο σύγκρινε τρία

---

<sup>2</sup> Το μοντέλο Hausman εκτιμάται με τη μέθοδο μεγίστης πιθανοφάνειας (ML)

<sup>3</sup> Οι περιορισμοί αυτοί έχουν να κάνουν με έναν μη γραμμικό εισοδηματικό περιορισμό (nonlinear-piecewise linear budget constraint) που αντιστοιχεί στις συνθήκες του Slutsky.

διαφορετικά είδη δεδομένων που αφορούσαν το ίδιο δείγμα ανθρώπων στη Σουηδία. Βασικός στόχος του ήταν να συγκρίνει μεθοδολογίες που μετρούν καλύτερα τις πραγματικές ώρες εργασίας (time-use studies), έναντι άλλων πιο συμβατικών μεθόδων που μετρούν τις κανονικές ώρες εργασίας που υπάρχουν στο συμβόλαιο (regular/conventional labor supply surveys). Η μέθοδος time-use θεωρείται πολύ ισχυρή καθώς περιλαμβάνει πολλές λεπτομερείς πληροφορίες μέσω των οποίων γίνεται πολύ καλή προσέγγιση των πραγματικών ωρών εργασίας. Οι πληροφορίες αφορούν εργασιακές δραστηριότητες αλλά και μη εργασιακές καθώς και ανάλυση του ελεύθερου χρόνου-σχόλης. Περιλαμβάνει λεπτομέρειες όπως: χρόνος μετακίνησης από το σπίτι στο χώρο εργασίας, χρόνος διαλείμματος, υπερωρίες, ώρες που αφιερώνονται σε δευτερεύοντες δουλειές, ασυνήθιστες απουσίες από τη δουλειά λόγω έκτακτων περιστατικών (πχ άρρωστα παιδιά), ταξίδια που αφορούν τη δουλειά κ.ά. Θεωρητικά λέγεται ότι μπορεί να μετρηθεί και ο χρόνος που αφιερώνεται στη 'μαύρη' αγορά.

Συγκρίνοντας τις κατανομές των εβδομαδιαίων ωρών εργασίας μεταξύ time-use και conventional δεδομένων για τη χρονιά 1993, που αντλήθηκαν από το Swedish household panel surveys-HUS, διαπίστωσε ότι τα time-use δεδομένα έχουν μεγαλύτερη διακύμανση και πιο ομαλή κατανομή. Οι εκτιμήσεις-συγκρίσεις αφορούν ημερήσιες και εβδομαδιαίες ώρες εργασίας (από time-use survey) και ετήσιες ώρες εργασίας (από regular survey). Οι εκτιμήσεις έγιναν με τη μέθοδο bootstrap 3SLS και το πρόβλημα επιλογής (*selection problem*) αντιμετωπίστηκε με τη προσέγγιση Heckit. Φυσικά για να γίνει η σύγκριση των τριών προσεγγίσεων των ωρών εργασίας, έγιναν οι κατάλληλες υπολογιστικές μετατροπές, αφού η κάθε προσέγγιση έχει διαφορετική μονάδα μέτρησης-χρονικό ορίζοντα (ημέρα, εβδομάδα, έτος). Έτσι, αναφορικά με τις ελαστικότητες μισθού, βρέθηκε ότι τα ετήσια δεδομένα των ωρών εργασίας δίνουν τις μικρότερες εκτιμήσεις, ενώ οι εβδομαδιαίες ώρες εργασίας δίνουν τις μεγαλύτερες και οι ημερήσιες ώρες είναι κάπου στη μέση. Αντίθετα, οι εισοδηματικές ελαστικότητες των ετήσιων ωρών εργασίας είναι σε απόλυτες τιμές οι μεγαλύτερες και για τα δύο φύλα. Σύμφωνα με τον συγγραφέα, ένας λόγος στον οποίο μπορεί να οφείλονται οι μικρές ελαστικότητες των ετήσιων ωρών εργασίας σε σχέση με το μισθό, είναι η μικρή μεταβλητότητα που υπάρχει στα δεδομένα των κανονικών ωρών εργασίας του συμβολαίου σε σχέση με τις άλλες δύο προσεγγίσεις.

H Birch (2005) έκανε μία ανασκόπηση των μελετών που έχουν γίνει στην Αυστραλία, με σκοπό να διαπιστώσει πόσο έχει αλλάξει ο τρόπος σκέψης και η μεθοδολογία που ακολουθείται στην εκτίμηση της προσφοράς εργασίας τα τελευταία 50 χρόνια. Οι αναφορές της δεν εστιάζονται μόνο στην περιοχή της Αυστραλίας, αλλά και σε έρευνες που έχουν πραγματοποιηθεί σε άλλες χώρες (όπως στις ΗΠΑ, Καναδά, Αγγλία κ.ά.), με σκοπό να

συγκρίνει τις λιγοστές αυστραλιανές μελέτες με τις πολυάριθμες μελέτες που έχουν λάβει χώρα σε άλλες ηπείρους. Η ανασκόπηση της βιβλιογραφίας εξετάζει τους παράγοντες που παίζουν ρόλο στην απόφαση που παίρνει μία γυναίκα για να μπει στην αγορά εργασίας και στο πόσες ώρες θα εργαστεί. Με λίγα λόγια, συγκεντρώνει όλες τις ερμηνευτικές μεταβλητές που έχουν αναφερθεί στην οικονομική θεωρία και έχουν χρησιμοποιηθεί κατά καιρούς στα οικονομετρικά μοντέλα και που πιστεύεται ότι πρέπει να λαμβάνονται υπόψη στην εξίσωση προσφοράς εργασίας.

Εν συντομία, οι παράγοντες που μπορούν να επηρεάσουν την εργασιακή συμπεριφορά των γυναικών είναι οικονομικοί (μισθός, εισόδημα, δάνεια, ποσοστό ανεργίας), προσωπικοί (μορφωτικό επίπεδο, επαγγελματική εμπειρία, οικογενειακή κατάσταση, αριθμός παιδιών, ηλικία παιδιών), δημογραφικοί (ηλικία, εθνικότητα, τόπος κατοικίας), και θεσμικοί-κοινωνικοί (επιδόματα, μεταβιβαστικές πληρωμές, ασφάλιση, φορολογικό σύστημα). Φυσικά δεν έχουν όλοι οι παράγοντες την ίδια βαρύτητα και εξαρτάται κάθε φορά από τη χώρα, το δείγμα και τα δεδομένα που έχει κάποιος. Υπό αυτή την έννοια είναι αυτονόητο ότι δεν μπορεί ένας ερευνητής να χρησιμοποιεί κάθε φορά όλες τις ανεξάρτητες μεταβλητές και ίσως η περίληψη τους να μην είναι απαραίτητη βάσει της στατιστικής σημαντικότητας. Για παράδειγμα η εθνικότητα φαίνεται να μην παίζει μεγάλο ρόλο στην Αυστραλία, καθώς οι μελέτες που έχουν γίνει δεν δίνουν ξεκάθαρα αποτελέσματα.

Αναφορικά τώρα με τη μεθοδολογία που χρησιμοποιήθηκε πριν πολλές δεκαετίες και τη μεθοδολογία που ακολουθείται τα τελευταία χρόνια, η συγγραφέας ταξινομεί τις μελέτες σε πρώτης και δεύτερης γενιάς (Killingsworth, 1983). Οι μελέτες πρώτης γενιάς εμφανίστηκαν τη δεκαετία του '60 και '70 και είναι αυτές που χρησιμοποιούσαν απλές και κλασικές μεθόδους εκτίμησης (όπως η OLS), με πολλά οικονομετρικά προβλήματα και εξειδίκευση των μοντέλων προσφοράς εργασίας που ερχόταν σε σύγκρουση με την οικονομική θεωρία. Αντίθετα οι μελέτες δεύτερης γενιάς που εμφανίστηκαν γύρω στο 1980 και κυριαρχούν μέχρι σήμερα, χρησιμοποιούν πιο εξελιγμένες και πολύπλοκες μεθόδους όπως είναι το Tobit και διάφορα μοντέλα δύο βημάτων (*two-step models*) για τη διόρθωση του sample selection bias. Στο πρώτο βήμα εκτιμάται η πιθανότητα συμμετοχής στην αγορά εργασίας, συγκρίνοντας το μισθό επιφύλαξης με το μισθό της αγοράς. Αυτό το βήμα εκτιμάται συνήθως με τα διωνυμικά μοντέλα Probit και Logit. Στη συνέχεια, στο δεύτερο βήμα, εκτιμάται ένα επαυξημένο μοντέλο προσφοράς εργασίας, καθώς στην εξίσωση των ωρών εργασίας συμπεριλαμβάνεται μία μεταβλητή από το πρώτο βήμα που διορθώνει το πρόβλημα της μεροληψίας. Αυτή η μεταβλητή εξασφαλίζει ότι ο διαταρακτικός όρος της εξίσωσης των ωρών εργασίας είναι μία τυχαία μεταβλητή με μέσο μηδέν, διαφορετικά δεν θα

είχε μηδενικό μέσο, καθώς θα συσχετιζόταν με τον διαταρακτικό όρο της εξίσωσης του πρώτου βήματος και κατ' επέκταση και με τις ερμηνευτικές μεταβλητές, δίνοντας έτσι μη συνεπείς εκτιμήσεις. Το δεύτερο βήμα εκτιμάται συνήθως με την απλή μέθοδο ελαχίστων τετραγώνων ή με τη γενικευμένη μέθοδο ελαχίστων τετραγώνων (OLS ή GLS). Πάντως οι περισσότερες αυστραλιανές μελέτες δίνουν βάρος στα μοντέλα δύο βημάτων για τη διόρθωση της μεροληψίας, ενώ αντίθετα λίγες είναι αυτές που χρησιμοποιούν το μοντέλο Tobit, αφού το Tobit δεν ελέγχει για μεροληψία δείγματος. Εξάιρεση αποτελεί η έρευνα των Kidd and Ferko (2001) στην οποία χρησιμοποιούνται και οι δύο τρόποι προσέγγισης.

Για ακόμη μία φορά, αυτό που τονίζει η συγγραφέας και είναι ευρέως διαδεδομένο στην υπάρχουσα βιβλιογραφία, είναι ότι υπάρχουν διαφορές στα μεγέθη των ελαστικοτήτων. Για παράδειγμα οι μελέτες πρώτης γενιάς δίνουν γενικά μικρότερες εκτιμήσεις ελαστικοτήτων σε σχέση με τις μελέτες δεύτερης γενιάς. Στην Αυστραλία αυτό ισχύει με τις εισοδηματικές ελαστικότητες, αλλά παρατηρείται το αντίθετο με τις ελαστικότητες συμμετοχής στην αγορά ως προς το μισθό, με τις μελέτες πρώτης γενιάς να δίνουν μία μέση ελαστικότητα συμμετοχής στην αγορά εργασίας γύρω στο 0,92 ενώ η αντίστοιχη ελαστικότητα δεύτερης γενιάς είναι κατά μέσο όρο 0,66. Βέβαια οι διαφορές αυτές μεταξύ των μελετών πρώτης και δεύτερης γενιάς δεν οφείλονται μόνο στις διαφορετικές μεθόδους εκτίμησης και την εξειδίκευση του μοντέλου αλλά και στα δεδομένα που ανήκουν σε διαφορετικές χρονικές περιόδους. Μία καλύτερη συγκριτική ανάλυση θα ήταν αν στα δεδομένα της δεκαετίας του '60 και '70 χρησιμοποιούσαμε οικονομετρικές μεθόδους δεύτερης γενιάς και έπειτα συγκρίναμε με τα αποτελέσματα της πρώτης γενιάς. Εκτός όμως από όλα αυτά που έχουν αναφερθεί και παραπάνω, ένα ακόμα στοιχείο που επηρεάζει τα μεγέθη των ελαστικοτήτων είναι το αν το δείγμα είναι ατομικό ή σε επίπεδο νοικοκυριού. Για παράδειγμα η μέση εισοδηματική ελαστικότητα προσφοράς εργασίας για τις γυναίκες, βάσει οικονομετρικών μοντέλων που στηρίζονται σε ατομικά δεδομένα, είναι -0,26. Η αντίστοιχη ελαστικότητα για τις γυναίκες, με βάση δεδομένα νοικοκυριού, είναι -1,48. Τέτοιες διαφορές όμως υπάρχουν και σε άλλες ελαστικότητες, όπως στην ελαστικότητα των ωρών εργασίας με βάση το μισθό. Γενικά οι περισσότερες μελέτες στην Αυστραλία συμφωνούν στο ότι οι εισοδηματικές ελαστικότητες των γυναικών έχουν αρνητικό πρόσημο. Υπάρχουν ελάχιστες μόνο που αντιβαίνουν τη θεωρία. Μία από αυτές είναι των Kidd & Ferko (2001) η οποία, σύμφωνα με τη Birch (2005), θεωρείται από τις πιο λεπτομερείς έρευνες που έχουν γίνει στην Αυστραλία.

Οι Kidd & Ferko (2001) ασχολούνται με τη διάκριση των φύλων και τις επιπτώσεις που έχει στους μισθούς και στην πιθανότητα εισόδου στην αγορά εργασίας. Σύμφωνα με



αυτούς, το επίπεδο του οικογενειακού εισοδήματος δεν επηρεάζει τη πιθανότητα εισόδου της γυναίκας στην αγορά εργασίας. Με λίγα λόγια, οι εισοδηματικές ελαστικότητες είναι μηδέν. Επίσης δίνεται έμφαση στην εξειδίκευση της εξίσωσης του μισθού επιφύλαξης στο πρώτο βήμα ενός μοντέλου δύο βημάτων, αφού φαίνεται να επηρεάζει σημαντικά τη πιθανότητα συμμετοχής στην αγορά εργασίας ανάλογα με το αν στην εξίσωση περιλαμβάνονται εναλλακτικές μεταβλητές όπως η εμπειρία, το τετράγωνο της εμπειρίας ή καμία από τις δύο.

Εν κατακλείδι τα συμπεράσματα που βγάζει η Birch (2005) για τις γυναίκες στην Αυστραλία, είναι ότι ο μισθός παίζει μεγαλύτερο ρόλο στην απόφαση να μουν στην αγορά εργασίας (ή και να βγουν) παρά στο πόσες ώρες θα εργαστούν. Στο ίδιο συμπέρασμα καταλήγει και ο Heckman (1993) με μία ανασκόπηση που είχε κάνει για τις γυναίκες στις ΗΠΑ. Επίσης διάφοροι εξωτερικοί παράγοντες της οικονομίας, όπως η αυξημένη ανεργία σε περιόδους χαμηλής οικονομικής δραστηριότητας, πιθανόν να δημιουργήσουν κάποια αρνητικά φαινόμενα σαν αυτό του επιπρόσθετου εργάτη<sup>4</sup> (*added-worker effect*). Επιπλέον, η παρουσία των παιδιών αλλά και η ηλικία τους φαίνεται να είναι καθοριστικής σημασίας, αφού παρατηρείται μία συνάρτηση προσφοράς εργασίας σε σχήμα 'm'. Δηλαδή η προσφορά εργασίας αυξάνεται μέχρι να φτάσουν οι γυναίκες στην αναπαραγωγική τους ηλικία, στη συνέχεια μειώνεται όταν αποκτήσουν παιδιά, αυξάνεται πάλι και φτάνει στο δεύτερο μέγιστο όταν τα παιδιά πάνε στο σχολείο, και μειώνεται συνεχώς όσο πλησιάζει η στιγμή να βγουν στη σύνταξη. Συγκρίνοντας τις μελέτες της Αυστραλίας με τις μελέτες άλλων χωρών, φαίνεται πως οι γυναίκες της Αυστραλίας μειώνουν τη προσφορά εργασίας τους περισσότερο από τις γυναίκες της Αγγλίας, του Καναδά και των ΗΠΑ όταν αυξάνεται η γονιμότητα. Με λίγα λόγια οι γυναίκες στην Αυστραλία είναι πιο ευαίσθητες στην παρουσία μικρών παιδιών απ' ότι οι αντίστοιχες άλλων χωρών. Τέλος η Birch (2005) θεωρεί ότι οι αυστραλιανές μελέτες έχουν μείνει λίγο πίσω σε σχέση με τις μελέτες του εξωτερικού, καθώς δεν περιλαμβάνουν στην έρευνά τους κάποιες μεταβλητές (όπως πχ την υγεία των γυναικών και τη φροντίδα ηλικιωμένων συγγενικών προσώπων) και δεν εξετάζουν τις μεταβολές στη προσφορά εργασίας των γυναικών κατά τη διάρκεια της ζωής τους (Heckman and MaCurdy, 1980).

Σύμφωνα με τα παραπάνω καταλαβαίνει κανείς ότι η παρουσία των παιδιών έχει καταλυτική σημασία στην προσφορά εργασίας, κυρίως των γυναικών. Είναι μία ιδιάζουσα παράμετρος που δεν λείπει σχεδόν ποτέ από ένα οικονομετρικό μοντέλο. Η προσέγγιση της

---

<sup>4</sup> Το φαινόμενο του επιπρόσθετου εργάτη παρατηρείται όταν το άτομο αυξάνει τη προσφορά εργασίας του, προκειμένου να αποζημιωθεί κατά κάποιον τρόπο από την επιδείνωση των οικονομικών συνθηκών (βλέπε Heckman and MaCurdy, 1980).

μεταβλητής αυτής μπορεί να γίνει εύκολα με τη χρήση ψευδομεταβλητών (0 αν δεν υπάρχει κάποιο παιδί - 1 αν υπάρχει, ή εναλλακτικά 0 αν δεν υπάρχει παιδί, 1 αν υπάρχει ένα παιδί, 2 αν υπάρχουν δύο, 3 αν υπάρχουν παραπάνω κ.α.). Εκτός όμως από αυτή τη μεταβλητή, υπάρχει και η ηλικία των παιδιών που ίσως είναι πιο σημαντική. Υπάρχουν ερευνητές που στα οικονομετρικά υποδείγματα χρησιμοποιούν μόνο μία εκ των δύο, αλλά το καλύτερο είναι να χρησιμοποιούνται και οι δύο μεταβλητές για να βγαίνουν και ορθότερα συμπεράσματα. Όσον αφορά την ηλικία, αυτό που συνήθως κάνουν οι ερευνητές είναι να κατηγοριοποιούν τη μεταβλητή της ηλικίας των παιδιών σε διαφορετικά ηλικιακά σύνολα, γιατί η κάθε κατηγορία επηρεάζει διαφορετικά τη προσφορά εργασίας των γυναικών. Φαίνεται πως τον μεγαλύτερο αντίκτυπο στη μείωση των ωρών εργασίας ή στη πιθανότητα εισόδου στην αγορά εργασίας έχουν τα νεογέννητα και τα μικρά παιδιά μέχρι 6 χρονών, ενώ συνήθως τα παιδιά άνω των 9 χρόνων δεν επηρεάζουν την εργασιακή συμπεριφορά των γυναικών. Επίσης υπάρχει περίπτωση τα παιδιά κάτω από 6 χρονών να μην επηρεάζουν τόσο πολύ τη προσφορά εργασίας των γυναικών, όταν υπάρχει και παρουσία μεγαλύτερων παιδιών στο σπίτι. Σε αυτή τη περίπτωση τα μεγαλύτερα παιδιά μπορούν και προσέχουν τα μικρότερα αδέρφια τους, δίνοντας έτσι τη δυνατότητα στη μητέρα να συμμετέχει πιο ενεργά στην αγορά εργασίας και να μην αναγκάζεται να μειώσει τις ώρες που εργάζεται.

Μία εκτενή ανάλυση στην ηλικία των παιδιών γίνεται από τους Leibowitz & Klerman (1995), οι οποίοι προσπαθούν να μελετήσουν τους παράγοντες που έχουν οδηγήσει στην αυξημένη προσφορά εργασίας των παντρεμένων μητέρων από το 1970 έως το 1990. Εκτός από τη παρουσία των παιδιών εξετάζουν και άλλους παράγοντες, κυρίως δημογραφικούς αλλά και οικονομικούς, που φαίνεται να έχουν συμβάλει σ' αυτή την αλλαγή. Από την ανάλυσή τους εξαιρούνται οι ανύπαντρες/ελεύθερες μητέρες, γιατί τις μεγαλύτερες αλλαγές τις εμφανίζουν οι παντρεμένες και επιπλέον θέλουν να δουν και την αλληλεπίδραση των συζύγων τους σ' αυτές τις αλλαγές. Το δείγμα αποτελείται από μητέρες με παιδιά από 1 έως και 36 μηνών και για να δουν λεπτομερώς την επίδραση των παιδιών, χωρίζουν τη μεταβλητή της παιδικής ηλικίας σε εννιά σύνολα. Σχεδόν όλα τα σύνολα δίνουν αρνητικές εκτιμήσεις και είναι στατιστικά σημαντικά. Αντίθετα, λαμβάνοντας υπόψη τη χρονική τάση, όλες οι εκτιμήσεις έχουν πλέον θετικό πρόσημο και είναι πολύ κοντά στο μηδέν και μόνο τρία σύνολα είναι στατιστικά σημαντικά (1, 2, και 3 μηνών). Επομένως φαίνεται πως τα νεογέννητα μωρά (1-3 μηνών) επηρεάζουν όλο και λιγότερο τη προσφορά εργασίας των παντρεμένων μητέρων μέσα στην εξεταζόμενη εικοσαετία. Γενικά όμως, αν εξαιρέσουμε την αλληλεπίδραση του χρόνου, τα συμπεράσματα είναι ίδια με αυτά που επικρατούν κατά

κανόνα παγκοσμίως. Δηλαδή η παρουσία πολύ μικρών παιδιών, κυρίως αυτών που δεν έχουν ξεκινήσει το δημοτικό σχολείο, μειώνουν τα ποσοστά συμμετοχής των γυναικών.

Επομένως, είδαμε σε αυτό το κεφάλαιο ότι οι εμπειρικές εκτιμήσεις των ελαστικότητων προσφοράς εργασίας διαφέρουν πολύ μεταξύ τους, ανάλογα με τις μεθόδους εκτίμησης, τη συναρτησιακή εξειδίκευση της εξίσωσης προσφοράς εργασίας, τα δεδομένα που έχει κάποιος στη διάθεσή του, τα λάθη μέτρησης στις ώρες εργασίας και το μισθό και πολλά άλλα. Η διαπίστωση αυτή έγινε και από τον Borjas (2003), ο οποίος αναφέρει χαρακτηριστικά ότι «υπάρχουν τόσοι τρόποι εκτίμησης της ελαστικότητας προσφοράς εργασίας, όσες είναι σχεδόν και οι εμπειρικές μελέτες στη βιβλιογραφία» (Borjas, 2003: σελ. 76).

## Κεφάλαιο 3

### Εναλλακτικές προσεγγίσεις στην εκτίμηση της συνάρτησης προσφοράς εργασίας

#### 3.1 Η διαχρονική (*lifetime*) συνάρτηση χρησιμότητας υπό το φάσμα του μοντέλου κύκλου-ζωής

Η έρευνα της Shaw (1992), περιορίζεται στη μελέτη της προσφοράς εργασίας μόνο των παντρεμένων γυναικών και εστιάζει τη προσοχή της κυρίως στο πώς επηρεάζεται η προσφορά εργασίας τους από το εισόδημα των συζύγων. Ισχυρίζεται μεταξύ των άλλων ότι το εισόδημα που παίζει μεγαλύτερο ρόλο είναι το μόνιμο-σταθερό (*permanent income*) και όχι το τρέχων εισόδημα (*current income*). Γι' αυτό το λόγο κάνει χρήση του μοντέλου του κύκλου ζωής<sup>5</sup> το οποίο υποδηλώνει αυτό ακριβώς το πράγμα, ότι δηλαδή σημασία δεν έχει τόσο το τρέχων εισόδημα, αλλά το εισόδημα που οι σύζυγοι προσδοκούν να έχουν μακροπρόθεσμα. Ο όρος *permanent income* δεν περιλαμβάνει μόνο τον προσδοκώμενο μισθό που προέρχεται από την εργασία, αλλά και όλα τα άλλα πιθανά έσοδα που δεν προέρχονται από αυτήν (όπως πχ μετοχές, ομόλογα κλπ). Έτσι το κάθε άτομο και με βάση την εκπαίδευση και εμπειρία που έχει αποκομίσει κατά τη διάρκεια της ζωής του, μπορεί να κάνει μία εκτίμηση του μόνιμου εισοδήματός του. Με λίγα λόγια η διαφορά του μόνιμου εισοδήματος από το τρέχων εισόδημα, είναι ότι το μόνιμο εισόδημα παίζει ουσιαστικά τον ρόλο του πλούτου<sup>6</sup>.

Σκοπός λοιπόν της συγγραφέως είναι να εκτιμήσει ένα life-cycle μοντέλο προσφοράς εργασίας και μέσω αυτού να βγάλει συμπεράσματα για την επιρροή του μόνιμου εισοδήματος των συζύγων και την εισοδηματική ανισότητα μεταξύ των νοικοκυριών. Τα δεδομένα τα αντλεί από το PSID (Panel Study of Income Dynamics) και αφορούν περίοδο δεκατεσσάρων ετών (1968-1981) για τις ΗΠΑ. Το δείγμα της αφορά μόνο παντρεμένες γυναίκες ηλικίας 18-64 που είναι τουλάχιστον τέσσερα χρόνια παντρεμένες με τον ίδιο σύζυγο. Έτσι συνολικά έχει ένα δείγμα με 19.103 παρατηρήσεις. Οι Donni & Moreau (2007) εστιάζουν επίσης σε ζευγάρια που έχουν τουλάχιστον δύο χρόνια κοινής συζυγικής ζωής. Ο λόγος που το κάνουν

---

<sup>5</sup> Το life-cycle model αναπτύχθηκε περίπου στα μέσα του 1950 από τους Modigliani & Brumberg (1954) και Milton Friedman (1957) με σκοπό τη πρόβλεψη της καταναλωτικής συμπεριφοράς των ανθρώπων διαχρονικά, λαμβάνοντας υπόψη το αναμενόμενο εισόδημα που θα αποκομίσουν κατά τη διάρκεια της ζωής τους. Έχει πολλές εφαρμογές, εξ' αυτών και στην αγορά εργασίας.

<sup>6</sup> Το μόνιμο εισόδημα υπολογίζεται ως  $\sum_{t=1}^T Y_t/T$  και το τρέχων εισόδημα ως απόκλιση από το μόνιμο εισόδημα:  $Y_t - \sum_{t=1}^T Y_t/T$

αυτό είναι γιατί οι κοινές αποφάσεις που παίρνει ένα νοικοκυριό (σχετικά με την εργασία του), είναι πιο σταθερές και χαρακτηρίζονται από μεγαλύτερη αποτελεσματικότητα όταν η οικογένεια είναι σχετικά αρκετό καιρό μαζί, σε σχέση με μια νεοσύστατη οικογένεια. Αυτή η αποτελεσματικότητα πηγάζει από το γεγονός ότι ένας άντρας και μία γυναίκα που είναι χρόνια παντρεμένοι, γνωρίζουν ο ένας τις προτιμήσεις του άλλου (Browning and Chiappori, 1998). Αντίθετα, ίσως να μην συμβαίνει το ίδιο με τα νεοσύστατα ζευγάρια. Αυτός πιθανόν να είναι και ο λόγος που και η συγγραφέας προτίμησε ένα δείγμα γυναικών που είναι χρόνια παντρεμένες.

Στόχος της Shaw (1992), είναι να συγκρίνει τις ελαστικότητες προσφοράς εργασίας τριών διαφορετικών μοντέλων: Tobit, Heckit και Probit. Το Probit εκτιμάει την πιθανότητα συμμετοχής των γυναικών στην εργασία, οπότε οι ελαστικότητες προσφοράς εργασίας αφορούν το extensive margin. Αντίθετα, οι ελαστικότητες που εκτιμώνται με το μοντέλο Heckit, αφορούν το intensive margin. Όμως, οι ελαστικότητες που εκτιμώνται με το μοντέλο Tobit αφορούν και το extensive και το intensive margin. Δηλαδή το Tobit λαμβάνει υπόψιν και τις δύο αποφάσεις που έχει να πάρει ο άνθρωπος σχετικά με μία θέση εργασίας (αν θα εισέλθει στην αγορά εργασίας και πόσες ώρες θα εργαστεί), με το να εκτιμάει μία συνολική ελαστικότητα προσφοράς εργασίας (Schultz, 1980; Shaw, 1992; Heckman, 1993; Birch, 2005). Ο Maglad (1998) αναφέρει ότι το άθροισμα της ελαστικότητας του probit με την ελαστικότητα του heckit προσεγγίζει καλύτερα την εκτίμηση του tobit, αν οι διαταρακτικοί όροι των εξισώσεων των δύο πρώτων μοντέλων είναι ανεξάρτητοι μεταξύ τους. Οι ελαστικότητες των γυναικών που εκτίμησε η Shaw (1992) αναφορικά με το μόνιμο εισόδημα του συζύγου είναι -0,51 με το Tobit, -0,19 με το Heckit και -0,38 με το Probit. Όπως φαίνεται, αν αθροίσουμε τις ελαστικότητες του heckit με του probit, προκύπτει η ελαστικότητα -0,57 που είναι κοντά στην ελαστικότητα του tobit. Οι αντίστοιχες ελαστικότητες αναφορικά με το τρέχον εισόδημα του συζύγου είναι -0,17, -0,07 και -0,11. Στη περίπτωση αυτή η συναθροιστική ελαστικότητα (0,07+0,11) είναι σχεδόν ίση με την ελαστικότητα του tobit (0,17).

Η βιβλιογραφία έχει δείξει ότι συνήθως οι ελαστικότητες προσφοράς εργασίας των γυναικών που εκτιμώνται με το Tobit είναι μεγαλύτερες (σε απόλυτες τιμές) από αυτές του Heckit (Schultz, 1980; Mroz, 1987). Στο ίδιο συμπέρασμα καταλήγει και η συγγραφέας. Επίσης, η βασική εξίσωση παλινδρόμησης που χρησιμοποιείται προϋποθέτει ότι δεν υπάρχει ενδογένεια μεταξύ του εισοδήματος του συζύγου και της προσφοράς εργασίας των γυναικών (Heckman and MaCurdy, 1980). Αντίθετα σε πολύ προσεχτικούς οικονομετρικούς ελέγχους που έχουν γίνει από άλλους ερευνητές (Mroz, 1987) έχει βρεθεί ότι υπάρχει ενδογένεια, αλλά

είναι τόσο μικρή που δεν επηρεάζει τα εμπειρικά μοντέλα. Έτσι η συγγραφέας συνεχίζει περαιτέρω και χωρίζει τη περίοδο μελέτης σε δύο ίσα μέρη και τις γυναίκες σε δύο ηλικιακές ομάδες (18-32, 33-64). Τα αποτελέσματα που βγάζει είναι ότι η εισοδηματική ελαστικότητα μειώνεται όσο περνάνε τα χρόνια και για τις νέες γυναίκες αλλά και για τις μεγαλύτερες σε ηλικία (πχ από -0,49 σε -0,29 για τις νέες). Επίσης τα δημογραφικά δεδομένα του PSID δείχνουν μία αύξηση του ρυθμού συμμετοχής των γυναικών στην αγορά εργασίας από το 1968 έως το 1981. Επομένως αυτή η αύξηση του ρυθμού συμμετοχής σε συνδυασμό με την μειούμενη αρνητική ελαστικότητα υποδηλώνουν μία μείωση της ευαισθησίας των γυναικών στο εισόδημα των συζύγων. Με λίγα λόγια αυτό σημαίνει ότι οι παντρεμένες γυναίκες γίνονται ολοένα και πιο (οικονομικά) ανεξάρτητες από τους συζύγους τους.

Τέλος, όσον αφορά την εισοδηματική ανισότητα μεταξύ οικογενειών, πολλοί ισχυρίζονται ότι η ανισοκατανομή αυξάνεται όσο περισσότερες γυναίκες εισέρχονται στην αγορά εργασίας, και αυτό γιατί πιστεύουν ότι άντρες και γυναίκες παντρεύονται κάποιον-α με 'ίδια προσόντα' (*assortative mating*). Αν ανατρέξουμε στην ιστορία μπορούμε να δούμε ότι ανέκαθεν από τα παλιά τα χρόνια υπήρχε μία τάση οι πλούσιοι να παντρεύονται με πλούσιες, οι γόνιμοι αριστοκρατικών οικογενειών έψαχναν να βρουν κάποιον 'εφάμιλλο' για να ενώσουν τις ζωές τους, ενώ άτομα φτωχά έσμιγαν με φτωχούς. Υπήρχαν δηλαδή ταξικές διαφορές οι οποίες έκαναν το χάσμα των οικογενειακών εισοδηματικών ανισοτήτων να είναι στα ύψη. Βέβαια σήμερα τείνουν να εξαλειφθούν τέτοιου είδους ταξικές διαφορές, αλλά δεν παύουν να υπάρχουν ορισμένοι ευκατάστατοι άνθρωποι με πλούσια επαγγελματική καριέρα που ψάχνουν να βρουν ένα σύζυγο με τα ίδια προσόντα. Επομένως όταν υπάρχει το φαινόμενο αυτό (*assortative mating*) και οι γυναίκες δεν είναι ευαίσθητες στο εισόδημα των συζύγων, η εισοδηματική ανισότητα μεταξύ οικογενειών θα μεγαλώνει όσο αυξάνεται η συμμετοχή των γυναικών στο εργατικό δυναμικό (αφού και τα δύο μέλη της οικογένειας θα έχουν το εισόδημά τους). Αντίθετα, όταν οι γυναίκες είναι ευαίσθητες σε μία μεταβολή του εισοδήματος του συζύγου, η ανισότητα θα μικραίνει καθώς μία αύξηση του συζυγικού εισοδήματος θα μειώσει πολύ τις ώρες εργασίας των γυναικών τους και μπορεί να βγουν και εντελώς από την αγορά εργασίας (μόνο το ένα μέλος της οικογένειας έχει εισόδημα-ο σύζυγος)<sup>7</sup>. Άρα έτσι θα υπήρχε μία ισορροπία στην κατανομή του οικογενειακού εισοδήματος.

---

<sup>7</sup> Υπενθυμίζω εδώ ότι όλα αυτά ισχύουν με δεδομένο ότι η εισοδηματική ελαστικότητα έχει αρνητικό πρόσημο (μία αύξηση του εισοδήματος των συζύγων έχει ως αποτέλεσμα τη μείωση προσφοράς εργασίας των γυναικών)

Για να μετρήσει την ανισότητα, η συγγραφέας χρησιμοποιεί το συντελεστή μεταβλητότητας, ο οποίος μέσω μίας συνάρτησης μπορεί να υπολογιστεί και ξεχωριστά για τη γυναίκα και τον σύζυγο με σκοπό να φανεί το μερίδιο της γυναίκας στην εισοδηματική ανισότητα. Τα συμπεράσματα της μέτρησης είναι ότι η οικογενειακή εισοδηματική ανισότητα έχει αυξηθεί σε μικρό βαθμό από την πρώτη περίοδο (1968-1974) στη δεύτερη (1975-1981). Αυτό συμβαδίζει με τις μειωμένες αρνητικές ελαστικότητες, (την μικρότερη ευαισθησία των γυναικών στο εισόδημα των συζύγων). Όμως παρόλο που οι ελαστικότητες έχουν μειωθεί πολύ, ειδικά στις νεαρότερες γυναίκες, η οικογενειακή ανισότητα έχει αυξηθεί ελάχιστα όπως είπαμε προηγουμένως και αυτό συμβαίνει γιατί υπάρχουν άλλοι παράγοντες που αντισταθμίζουν την ανισότητα, όπως είναι η φθίνουσα διακύμανση των χρηματικών απολαβών των γυναικών.

### **3.2 Μεγιστοποίηση συνάρτησης χρησιμότητας υπό περιορισμούς και εμπειρική εκτίμηση**

Οι Aaberge et al. (1999) πραγματοποίησαν μία καινοτομική έρευνα που αφορά την Ιταλία, καθώς είναι οι πρώτοι που κατασκεύασαν ένα μοντέλο το οποίο περιλαμβάνει από κοινού τις προτιμήσεις των νοικοκυριών ως προς την αγορά εργασίας, τους περιορισμούς που θέτονται από την αγορά στην κατανομή των ωρών εργασίας και τους φορολογικούς συντελεστές εισοδήματος. Τα δεδομένα αντλήθηκαν από το SHIW (Survey of Household Income and Wealth), μία έρευνα που διεξήχθη από τη Κεντρική Τράπεζα της Ιταλίας και αφορούν το έτος 1987. Η κατασκευή του μοντέλου ξεκινά με την υπόθεση ότι το κάθε άτομο προσπαθεί να μεγιστοποιήσει την χρησιμότητα που απολαμβάνει από την εργασία του υπό κάποιους περιορισμούς. Υποθέτουν ότι το περιβάλλον του κάθε ατόμου αποτελείται από μία πληθώρα εργασιακών και μη εργασιακών ευκαιριών, με καθεμία δουλειά να έχει καθορισμένες ώρες εργασίας, επίπεδο μισθών καθώς και άλλα χαρακτηριστικά που αφορούν το χώρο εργασίας. Η χρησιμότητα επομένως είναι συνάρτηση του μισθού, των ωρών εργασίας και του εισοδήματος που δεν προέρχεται από την εργασία. Ο περιορισμός που αντιμετωπίζει το κάθε άτομο είναι ο φόρος εισοδήματος ή εναλλακτικά, το διαθέσιμο εισόδημα που του απομένει μετά την αφαίρεση των φόρων. Στη συνέχεια οι συγγραφείς πραγματοποιούν διεξοδικές αναλύσεις για να καταλήξουν στο τελικό μοντέλο: η συνάρτηση προσφοράς εργασίας θεωρείται συνεχής γι' αυτό και όλη η ανάλυση περιέχει συναρτήσεις πυκνότητας, υπό συνθήκη πιθανότητες, ολοκληρώματα κλπ, άσχετα με το αν οι συγγραφείς παρουσιάζουν μία διακριτή (και όχι συνεχή) εκδοχή του μοντέλου για λόγους ευκολίας.

Η έρευνα περιλαμβάνει άντρες και γυναίκες ηλικίας 20-68 χρόνων και συνολικά 2953 νοικοκυριά. Τα νοικοκυριά που το μεγαλύτερο μέρος του εισοδήματός τους προέρχεται από την αυτοαπασχόληση (συγκεκριμένα πάνω από το 20% του ακαθάριστου οικογενειακού εισοδήματος) έχουν αποκλειστεί από την έρευνα, καθώς οι ώρες εργασίας είναι μη παρατηρήσιμες και επιπλέον οι αποφάσεις που παίρνει ένας αυτοαπασχολούμενος χαρακτηρίζονται από μία διαρκή αβεβαιότητα και διαφέρουν από αυτές των μισθωτών, με αποτέλεσμα ίσως να αντιμετωπίζει δομικά μία διαφορετική συνάρτηση χρησιμότητας από τους μισθωτούς.

Η μεθοδολογία που ακολουθείται από τους συγγραφείς στην εκτίμηση των ελαστικότητων διαφέρει από άλλες μελέτες, καθώς οι περισσότερες εκτιμάνε την ελαστικότητα με βάση τη μέση τιμή της μεταβλητής του δείγματος, ενώ εδώ εκτιμώνται οι ακριβείς συναθροιστικές ελαστικότητες για όλο το δείγμα. Με αυτόν τον τρόπο ελέγχουν την ετερογένεια στις προτιμήσεις του πληθυσμού και μπορούν να βγάλουν πιο σαφή συμπεράσματα και γενικεύσεις για όλο το πληθυσμό. Έτσι εν συνεχεία υπολογίζουν με τη μέθοδο της προσομοίωσης τις ελαστικότητες και παρουσιάζουν τη μέση ελαστικότητα για κάθε κατηγορία. Οι κατηγορίες είναι τρεις και αφορούν την ελαστικότητα αναφορικά με: τη πιθανότητα συμμετοχής (probability of participation), την κανονική (unconditional) και την υπό συνθήκη (conditional) προσδοκία για τις προσφερόμενες ώρες εργασίας. Βέβαια τη κάθε κατηγορία την υπολογίζουν και σε υποδείγματα. Πχ χωρίζουν το δείγμα σε άντρες-γυναίκες και σε τρία οικονομικά στρώματα με βάση το διαθέσιμο εισόδημα (το 10% των πλουσιότερων νοικοκυριών, το 10% των φτωχότερων, και το 80% της μεσαίας τάξης). Όλες αυτές οι προσομοιώσεις έχουν γίνει με βάση τη δημογραφική και εκπαιδευτική δομή, το φορολογικό σύστημα και τους θεσμικούς περιορισμούς που επικρατούσαν εκείνη την εποχή στην Ιταλία. Επομένως μία αλλαγή σε κάποιον από αυτούς τους παράγοντες θα μεταβάλλει και τις ελαστικότητες.

Τα αποτελέσματα δείχνουν ότι οι ελαστικότητες των γυναικών είναι αρκετά μεγαλύτερες από των αντρών, ειδικά όσον αφορά τη πιθανότητα συμμετοχής. Ενδεικτικά, η ελαστικότητα συμμετοχής ως προς το μισθό είναι περίπου 0,05 για τους άντρες και 0,65 για τις γυναίκες. Αναφορικά με τα εισοδηματικά στρώματα, οι γυναίκες που ανήκουν στα φτωχότερα εμφανίζουν τις μεγαλύτερες ελαστικότητες. Συγκεκριμένα η compensated ελαστικότητα συμμετοχής ως προς το μισθό είναι 9,62 και η compensated ελαστικότητα ωρών εργασίας ως προς το μισθό είναι 12,1. Οι αντίστοιχες ελαστικότητες για τα μεσαία στρώματα είναι 0,65 και 0,76. Κάτι αξιοσημείωτο επίσης εδώ είναι ότι όλες οι ελαστικότητες ως προς το μισθό έχουν θετικό πρόσημο και για τους άντρες και για τις γυναίκες σε όλα τα



εισοδηματικά στρώματα, εκτός από τις uncompensated ελαστικότητες των αντρών των υψηλότερων εισοδημάτων. Αυτό σημαίνει ότι οι Ιταλοί με υψηλό διαθέσιμο εισόδημα θα μειώσουν τις ώρες εργασίας σε μία ενδεχόμενη αύξηση του μισθού τους. Με λίγα λόγια μπορούμε να ισχυριστούμε ότι σε αυτή τη κατηγορία αντρών υπερισχύει το αποτέλεσμα εισοδήματος, ενώ για όλους τους άλλους (άντρες και γυναίκες) επικρατεί το αποτέλεσμα υποκατάστασης. Αυτή είναι πολύ σημαντική λεπτομέρεια που πρέπει να ληφθεί υπόψη από το κράτος σε μία εφαρμογή (φορολογικής) πολιτικής, αλλά και από τις επιχειρήσεις. Για παράδειγμα μία επιχείρηση η οποία γνωρίζει ότι κάποιοι από τους απασχολούμενους έχουν υψηλά εισοδήματα, μπορεί αναλογικά να προσαρμόσει το ύψος των μισθών που τους προσφέρει.

Τέλος οι συγγραφείς προχωρούν σε προσομοιώσεις του φορολογικού συστήματος της Ιταλίας με σκοπό να δουν πώς επηρεάζεται η προσφορά εργασίας και άλλες μεταβλητές σε μία πιθανή μεταρρύθμιση. Η σύγκριση γίνεται μεταξύ τριών φορολογικών συστημάτων: του πραγματικού συστήματος που υπήρχε το 1987, ενός υποθετικού αναλογικού συστήματος φόρων και ενός υποθετικά υψηλού προοδευτικού συστήματος. Σύμφωνα με τα αποτελέσματα, οι αλλαγές που επιφέρουν οι μεταρρυθμίσεις των υποθετικών φορολογικών συστημάτων είναι πολύ μικρές. Συγκεκριμένα οι αναλογικοί φόροι αυξάνουν τις ώρες εργασίας των αντρών, ενώ αντίθετα μειώνεται η προσφορά εργασίας των γυναικών. Το ίδιο συμβαίνει και με τα έσοδα για κάθε φύλο. Όσον αφορά από κοινού το νοικοκυριό, το διαθέσιμο εισόδημά του αυξάνεται. Οι προοδευτικοί φόροι δεν επηρεάζουν σχεδόν καθόλου τη προσφορά εργασίας των αντρών, ενώ αντίθετα μειώνουν τη προσφορά των γυναικών και μάλιστα σε μεγαλύτερο βαθμό απ' ό,τι οι αναλογικοί. Επίσης το διαθέσιμο εισόδημα του νοικοκυριού μειώνεται.

Οι συγγραφείς θέλουν επιπλέον να δουν τι επίπτωση έχει μία μεταρρύθμιση του φορολογικού συστήματος στην ανισοκατανομή του ακαθάριστου και του διαθέσιμου εισοδήματος μεταξύ των νοικοκυριών, γι' αυτό και εκτιμούν το συντελεστή Gini<sup>8</sup>. Με το φορολογικό σύστημα του 1987, ο Gini για το ακαθάριστο εισόδημα είναι 0,253 και για το διαθέσιμο εισόδημα 0,238. Με ένα προοδευτικό σύστημα οι αντίστοιχοι συντελεστές είναι 0,255 και 0,220 ενώ η υιοθέτηση αναλογικών φόρων δίνει ένα Gini 0,247 και για τα δύο εισοδήματα. Επομένως βλέπουμε ότι η υιοθέτηση αναλογικών φόρων αυξάνει την ανισότητα κατά 3,8% ενώ το προοδευτικό σύστημα μειώνει την ανισότητα κατά 7,6% (αναφορικά με το

---

<sup>8</sup> Ο συντελεστής Gini είναι ένας δείκτης μέτρησης της ανισότητας και παίρνει τιμές από 0 έως 1, με το 0 να υποδεικνύει την τέλεια ισότητα και το 1 την τέλεια ανισότητα.

διαθέσιμο εισόδημα). Άρα ένα ισχυρό προοδευτικό σύστημα φαίνεται να είναι η καλύτερη λύση έτσι ώστε να υπάρχει μεγαλύτερη ισοκατανομή του εισοδήματος στην Ιταλία.

H Dessing (2002), όπως αναφέρθηκε ήδη στο κεφάλαιο 1, εστιάζεται στις φτωχές οικογένειες και στη στρατηγική συμπεριφορά των μελών της για διαφορετικά επίπεδα μισθού. Τα στοιχεία που χρησιμοποιεί η συγγραφέας είναι από μία έρευνα που έγινε σε μία φτωχή αγροτική περιοχή στις Φιλιππίνες μεταξύ 1975 και 1976 από εξειδικευμένες ομάδες που παρακολουθούσαν από κοντά τους συμμετέχοντες και έκαναν συνολικά τέσσερις κατ' οίκων επισκέψεις σε κάθε νοικοκυριό μέσα σε διάστημα 8 μηνών. Η έρευνα περιλαμβάνει συνολικά 99 νοικοκυριά και τα δεδομένα είναι χρονικά καταναμημένα. Το μέγεθος του δείγματος μπορεί να θεωρείται σχετικά μικρό για να βγάλει κανείς σίγουρα συμπεράσματα, αλλά όπως περιγράφει η συγγραφέας, η μέθοδος που ακολουθήθηκε από τους ερευνητές είναι από τις πιο αξιόπιστες και συνεπείς που έχουν γίνει στη συγκεκριμένη περιοχή. Φυσικά η έρευνα περιλαμβάνει και βασικά στοιχεία για κάθε άτομο ξεχωριστά (πχ ποιος είναι ο βασικός εργάτης της οικογένειας, έτη εκπαίδευσης κάθε μέλους, ηλικία, ατομικός μισθός και ώρες εργασίας σε περίπτωση που εργάζονται και οι σύζυγοι και τα παιδιά κ.α.). Όμως η συγγραφέας εκτός από το να αναδείξει τις ατέλειες του νεοκλασικού μοντέλου, θέλει επίσης να παρατηρήσει τις αλληλοεξαρτώμενες αποφάσεις που παίρνουν τα μέλη μιας οικογένειας αναφορικά με την εργασία τους. Έτσι επεκτείνει το μοντέλο προσφοράς εργασίας σε οικογενειακό επίπεδο (επίπεδο νοικοκυριού).

Η οικογένεια λοιπόν προσπαθεί να μεγιστοποιήσει τη χρησιμότητά της (η χρησιμότητα είναι συνάρτηση τηςσχόλης και της κατανάλωσης) υπό τρεις περιορισμούς: τον χρονικό περιορισμό, τον εισοδηματικό περιορισμό και τον περιορισμό διαβίωσης. Ο χρονικός περιορισμός έχει να κάνει πολύ απλά με τον μέγιστο αριθμό διαθέσιμων ωρών που έχει η οικογένεια (πχ μέσα σε μια ημέρα) και οι οποίες αναλώνονται σε ώρες εργασίας και σε ώρεςσχόλης<sup>9</sup>. Ο εισοδηματικός περιορισμός είναι ο συνήθης περιορισμός που λέει ότι όλο το εισόδημα πηγαίνει στη κατανάλωση και ο περιορισμός διαβίωσης, όπως περιγράφηκε και προηγουμένως, υποδηλώνει το ελάχιστο εισόδημα που πρέπει να κερδίσει η οικογένεια για να καλύψει τις βασικές καταναλωτικές της ανάγκες. Στο σημείο που βρίσκεται ο μισθός διαβίωσης (*subsistence wage*) η οικογένεια αφιερώνει όλο το διαθέσιμο χρόνο της για εργασία, δηλαδή δεν έχει καθόλου ελεύθερο χρόνο. Στο σημείο αυτό δουλεύουν φυσικά όλα τα μέλη της οικογένειας, ακόμα και τα μικρά παιδιά. Ο μισθός διαβίωσης μπορεί να παρομοιαστεί με το μισθό επιφύλαξης του νεοκλασικού μοντέλου, αλλά εννοιολογικά δεν

---

<sup>9</sup> Αναφέρω εδώ ότι στις ώρεςσχόλης, δηλαδή στον ελεύθερο χρόνο, δεν συμπεριλαμβάνονται οι ώρες ύπνου

πρέπει να συγχέονται καθώς ο μισθός επιφύλαξης έγκειται καθαρά σε υποκειμενικούς λόγους κάθε ατόμου, ενώ ο μισθός διαβίωσης ‘επιβάλλεται’ κατά κάποιο τρόπο στο άτομο από φυσικούς παράγοντες.

Αυτό που συνήθως παρατηρείται στις αγροτικές οικογένειες είναι ότι εφόσον έχει εξασφαλιστεί το αναγκαίο εισόδημα, τότε όσο αυξάνεται ο μισθός, τόσο λιγότερες ώρες δουλεύουν οι γυναίκες και τα παιδιά. Αυτό δηλαδή που υποδηλώνει η συγγραφέας είναι ότι η εργασία των δευτερευόντων μελών της οικογένειας είναι συμπληρωματική/βοηθητική σε περίπτωση που δεν μπορεί να τα βγάλει πέρα μόνος του ο βασικός εργάτης (που συνήθως είναι ο άντρας). Έτσι τα δευτερεύοντα μέλη εργάζονται μόνο στα πολύ χαμηλά και στα υψηλά επίπεδα μισθού (από τον μισθό επιφύλαξης και πάνω), ενώ στο μεσαίο κομμάτι αποχωρούν από την αγορά εργασίας. Ο λόγος που οι γυναίκες μειώνουν τις ώρες εργασίας τους εφόσον έχουν καλυφθεί οι βασικές ανάγκες, είναι γιατί πρέπει να ασχοληθούν και με τις δουλειές στο σπίτι, κυρίως με τη φροντίδα των παιδιών. Θα σταματήσουν εντελώς να εργάζονται μόνο όταν πλέον δεν χρειάζεται καμία βοήθεια ο άντρας και τότε αφιερώνουν όλο το χρόνο τους στην ανατροφή των παιδιών, στα οικιακά και σεσχόλη. Από αυτό το σημείο και μετά δρουν έτσι όπως ακριβώς ‘προστάζει’ το νεοκλασικό μοντέλο: δηλαδή θα ξαναμπούν στην αγορά εργασίας μόνο αφού συναντήσουν τον μισθό επιφύλαξης.

Επιστρέφοντας λοιπόν πίσω στο καθαρά εμπειρικό κομμάτι της συγγραφής, το ομοιογενές και μη περικομμένο (*non-truncated*) δείγμα των νοικοκυριών μπορεί να είναι μικρό σε μέγεθος αλλά αντισταθμίζεται πλήρως από την ακρίβεια των μετρήσεων των μεταβλητών. Η εξίσωση παλινδρόμησης έχει την εξής απλή μορφή:  $\ln L = \alpha + \beta \ln w + \delta X + \theta$ , όπου  $L$  οι ώρες εργασίας που αφιερώνει η οικογένεια σε όλες τις δραστηριότητές της (πχ αγροτική παραγωγή, εγχώρια παραγωγή κλπ),  $w$  ο μισθός, και  $X$  ένα διάνυσμα μεταβλητών που αφορούν ατομικά και οικογενειακά χαρακτηριστικά, με το  $\theta$  να είναι το τυχαίο σφάλμα. Όπως γίνεται εύκολα αντιληπτό από τα μέχρι τώρα συμφραζόμενα, η εκτίμηση της εξίσωσης παλινδρόμησης αφορά μόνο οικογένειες με χαμηλά εισοδήματα. Εν ολίγοις οι εκτιμημένες ελαστικότητες αφορούν το τμήμα *low-subsistence wage* της *S*-καμπύλης προσφοράς εργασίας (Εικόνα 1c). Η εξίσωση εκτιμάται γενικά για όλη την οικογένεια αλλά και ξεχωριστά για τους βασικούς και δευτερεύοντες εργάτες. Τα εμπειρικά αποτελέσματα φαίνεται να είναι συνεπή με τις θεωρητικές υποθέσεις που έχουν διατυπωθεί παραπάνω, αναφορικά με τη στρατηγική συμπεριφορά των φτωχών οικογενειών. Συγκεκριμένα, η ελαστικότητα προσφοράς εργασίας είναι πολύ κοντά στο μηδέν (0,13) για τους βασικούς εργάτες, ενώ για τους δευτερεύοντες εργάτες και γενικά για την οικογένεια είναι μεγαλύτερες και αρνητικές (-0,39 και -0,55 αντίστοιχα). Όμως μόνο οι δύο αυτές τελευταίες ελαστικότητες

είναι στατιστικά σημαντικές. Η σχεδόν μηδενική και στατιστικά ασήμαντη ελαστικότητα του βασικού εργάτη, μπορούμε να πούμε ότι αντικατοπτρίζει το γεγονός που έχει αναφερθεί προηγουμένως στο θεωρητικό υπόβαθρο: ότι σαν πρωταρχικός εργάτης που είναι, εργάζεται όσο περισσότερο μπορεί προκειμένου να συντηρήσει την οικογένειά του και οι ώρες εργασίας του δεν επηρεάζονται από το επίπεδο του μισθού, αλλά από άλλους (φυσικούς) παράγοντες. Επιπλέον οι δύο στατιστικά σημαντικές ελαστικότητες έχουν αρνητικό πρόσημο, γεγονός που επιβεβαιώνει την κριτική της συγγραφέως στο νεοκλασικό μοντέλο περί θετικών ελαστικοτήτων. Αυτά τα αποτελέσματα εκτιμήθηκαν με τη μέθοδο ελαχίστων τετραγώνων.

Επίσης η συγγραφέας τρέχει και το υπόδειγμα του Heckman για τους δευτερεύοντες εργάτες προκειμένου να ελέγξει για μεροληψία από την επιλογή του δείγματος (*sample selection bias*). Η εκτιμημένη ελαστικότητα έχει σχεδόν την ίδια τιμή με την OLS ελαστικότητα και το *inverse Mill's ratio* δεν είναι στατιστικά σημαντικό, υποδεικνύοντας έτσι τη συνέπεια των αρχικών ευρημάτων. Αντίθετα, σύμφωνα με την ομάδα που έκαναν την έρευνα στις Φιλιππίνες, ίσως να υπάρχει πρόβλημα με την μέτρηση του πλούτου της οικογένειας. Γι' αυτό το λόγο, η συγγραφέας που χρησιμοποιεί τον πλούτο ως ανεξάρτητη μεταβλητή, προχωράει και σε εκτίμηση με τη μέθοδο IV (*Instrumental Variable*) βάζοντας ως εργαλείο και τον πλούτο και το μισθό. Τα αποτελέσματα διαφέρουν από τη μέθοδο OLS μόνο ως προς την οικογένεια, όπου η ελαστικότητα από αρνητική έχει γίνει οριακά θετική (0,06). Τα αποτελέσματα για τους βασικούς εργάτες είναι σχεδόν ίδια (0,16) ενώ για τους δευτερεύοντες εργάτες έχει μειωθεί σε απόλυτες τιμές η ελαστικότητα αλλά παραμένει αρνητική (-0,16).

Κάνοντας μία σύνοψη και για τις δύο μεθόδους εκτίμησης (OLS και IV), βλέπουμε ότι η ελαστικότητα των γυναικών έχει αρνητικό πρόσημο που συμφωνεί με την υπόθεση ότι όσο αυξάνεται ο μισθός, η συμβολή τους στην εργασία μειώνεται με σκοπό να αφιερώσουν περισσότερο χρόνο στα παιδιά και στη γενικότερη ευημερία της οικογένειάς τους, ενώ η θετική ελαστικότητα των αντρών υποδηλώνει ότι σαν πρωταρχικοί εργάτες θα αυξάνουν τις ώρες εργασίας τους συνεχώς, έως ότου φτάσουν στο μέγιστο αριθμό ωρών που μπορούν να αντέξουν (σωματικά και ψυχικά) προκειμένου να επιβιώσει η οικογένειά τους. Η εναλλαγή προσήμων στην ελαστικότητα της οικογένειας μπορεί να οφείλεται στα λάθη μέτρησης του πλούτου του κάθε νοικοκυριού, με συνέπεια να μην είναι ευδιάκριτο αν επικρατεί το αποτέλεσμα εισοδήματος ή υποκατάστασης. Πάντως φαίνεται να είναι ισχυρότερο το αποτέλεσμα εισοδήματος αν κρίνουμε συναθροιστικά την αρνητική και σχεδόν μηδενική εκτίμηση που έδωσαν η μέθοδος OLS και IV αντίστοιχα, και αν θυμηθούμε ότι η S-καμπύλη προσφοράς εργασίας στα χαμηλά επίπεδα μισθού έχει αρνητική κλίση.

Σ' αυτό το κεφάλαιο είδαμε με λίγα λόγια ότι η συνάρτηση προσφοράς εργασίας μπορεί να εκτιμηθεί υπό διαφορετικούς περιορισμούς. Οι βασικοί περιορισμοί περιλαμβάνουν συνήθως τον κλασικό εισοδηματικό περιορισμό και τον χρονικό περιορισμό. Ανάλογα όμως με τη φύση της μελέτης, μπορεί να περιλαμβάνει και άλλους περιορισμούς όπως είναι ο περιορισμός διαβίωσης (Dessing, 2002). Επίσης είδαμε ότι η προσφορά εργασίας μπορεί να μελετηθεί με ένα δυναμικό υπόδειγμα, μέσω πχ ενός μοντέλου κύκλου ζωής, αλλά και με στατικά υποδείγματα, τα οποία εμφανίζονται συχνότερα στη βιβλιογραφία λόγω μεγαλύτερης ευκολίας στη συλλογή δεδομένων.

## Κεφάλαιο 4

### Το πρόβλημα της μεροληψίας από την επιλογή του δείγματος (sample selection bias)

Από αυτά που έχουν αναφερθεί μέχρι τώρα στα προηγούμενα κεφάλαια, μπορεί να πάρει κάποιος μία ιδέα για το ποιες είναι οι δυσκολίες που συναντάει ένας ερευνητής στη προσπάθειά του να εκτιμήσει εμπειρικά την προσφορά εργασίας και τις αντίστοιχες ελαστικότητες. Υπάρχουν αντικρουόμενες απόψεις μεταξύ των οικονομολόγων που ασχολούνται με την αγορά εργασίας (*labor economists*) ως προς το ποια είναι η ενδεδειγμένη μεθοδολογία που πρέπει να ακολουθεί κανείς στην εκτίμηση. Μέχρι τώρα δεν υπάρχει κάποια σαφή επιχειρηματολογία που να αποδεικνύει αν υπάρχει ένας καθορισμένος τρόπος εφαρμογής των οικονομετρικών εργαλείων, γι' αυτό το λόγο άλλωστε υπάρχουν μεγάλες αποκλίσεις στις εκτιμημένες ελαστικότητες και ποικίλοι τρόποι εκτίμησης της προσφοράς εργασίας στην βιβλιογραφία (όπως αναφέρθηκε σε προηγούμενα κεφάλαια). Κάποιοι ερευνητές θεωρούν ότι πρέπει να βρεθεί ένα πιο κοινό μοτίβο στην εξειδίκευση των μοντέλων και στη συλλογή δεδομένων, καθώς το εύρος των εκτιμήσεων δεν έχει μειωθεί σχεδόν καθόλου, παρόλο που πλέον υπάρχουν περισσότερες γνώσεις και χρησιμοποιούνται πιο εκλεπτυσμένα υποδείγματα σε σχέση με το παρελθόν (Birch, 2005). Παρακάτω παρουσιάζεται εκτενέστερα το σημαντικότερο ίσως πρόβλημα που έχει προκύψει στα οικονομικά της εργασίας και εναλλακτικοί τρόποι ελέγχου που έχουν προταθεί.

#### 4.1 Τι είναι η μεροληψία επιλογής δείγματος και που οφείλεται;

Είναι γνωστό ότι στις εμπειρικές μελέτες τα δείγματα πρέπει να συλλέγονται τυχαία από έναν εξεταζόμενο πληθυσμό, με σκοπό τα στατιστικά και οικονομετρικά αποτελέσματα να είναι όσο το δυνατόν περισσότερο αντιπροσωπευτικά και να γίνεται σωστή επαγωγή στον πληθυσμό (Χαρίσης και Κιόχος, 1997: σελ.20). Διαφορετικά το δείγμα είναι μεροληπτικό. Εκτός αυτού, οι περισσότερες οικονομετρικές τεχνικές και στατιστικοί έλεγχοι που έχουν αναπτυχθεί υποθέτουν ότι το δείγμα είναι τυχαίο. Επομένως οι τεχνικές αυτές δεν μπορούν να χρησιμοποιηθούν στην ανάλυση της προσφοράς εργασίας ενός δείγματος που περιλαμβάνει μόνο εργαζόμενους (Borjas, 2003: σελ. 79). Για να είναι τυχαίο το δείγμα θα πρέπει να περιλαμβάνει και τους άνεργους.

Στον κλάδο των οικονομικών της εργασίας το πρόβλημα του μη τυχαίου δείγματος είναι αρκετά σύννηθες, αφού οι ώρες εργασίας και ο μισθός που αποτελούν αναπόσπαστο

κομμάτι στην εκτίμηση ελαστικότητας, δεν μπορούν να παρατηρηθούν για ολόκληρο το πληθυσμό, παρά μόνο γι' αυτούς που εργάζονται. Έτσι συχνά το δείγμα αποτελείται μόνο από εργαζόμενους, θεωρώντας ότι για τους άνεργους οι ώρες εργασίας και ο μισθός έχουν μηδενικές τιμές και γι' αυτό εξαιρούνται από την ανάλυση. Στη πραγματικότητα όμως, ο μισθός ενός ατόμου που δεν εργάζεται δεν είναι μηδενικός, αλλά είναι ένα ποσό μικρότερο από τον μισθό επιφύλαξής του (Borjas, 2003: σελ. 79). Επομένως θα πρέπει να υπολογίζεται ο μισθός για τους άνεργους και να συμπεριλαμβάνονται και αυτοί στο δείγμα. Η μεροληψία που προκύπτει στο δείγμα όταν δεν περιλαμβάνει τους άνεργους είναι γνωστή ως μεροληψία επιλογής δείγματος (*sample selection bias*<sup>10</sup>). Αν εξαιρέσουμε τις ώρες εργασίας και το μισθό των ανέργων, όλες οι παρατηρήσεις που αφορούν τις υπόλοιπες συνήθεις μεταβλητές που χρησιμοποιούνται σε ένα μοντέλο προσφοράς εργασίας (πχ ηλικία, μορφωτικό επίπεδο, εμπειρία, οικογενειακή κατάσταση κλπ) είναι γνωστές. Έτσι προκύπτει ένα λογοκριμένο (*censored*) ή και περικομμένο (*truncated*<sup>11</sup>) δείγμα, το οποίο αν αντιμετωπιστεί με τη γραμμική μέθοδο των ελαχίστων τετραγώνων θα δώσει μεροληπτικές και μη συνεπείς εκτιμήσεις. Η σωστή αντιμετώπιση τέτοιων δειγμάτων γίνεται με το μοντέλο Tobit.

Κάποιοι ερευνητές όμως αποφεύγουν να υπολογίζουν τον μισθό των ανέργων και προτιμούν εξ' αρχής να ασχοληθούν με ένα δείγμα το οποίο θα περιλαμβάνει μόνο άτομα που εργάζονται, 'πετώντας' ουσιαστικά έξω από το δείγμα τα άτομα για τα οποία δεν έχουν όλες τις παρατηρήσεις (Borjas, 2003: σελ. 79; Donni and Moreau, 2007). Έτσι αυτόματα το δείγμα τους γίνεται μεροληπτικό-μη τυχαίο. Εναλλακτικά, αντί να μειώσουν το δείγμα τους, αυτό που κάνουν οι περισσότεροι είναι να υπολογίζουν τη μέση τιμή της λογοκριμένης μεταβλητής από όσους έχουν απαντήσει αυτή τη μεταβλητή και να τη χρησιμοποιούν ως παρατηρούμενη τιμή για αυτούς που δεν απάντησαν. Η διαδικασία αυτή υποθέτει ότι η μη ανταπόκριση του συνεντευξιαζόμενου σε μία ερώτηση του ερευνητή είναι τυχαία. Σε αυτή τη σπάνια περίπτωση, η διαδικασία της μέσης τιμής είναι αποδεκτή. Η πιο ρεαλιστική περίπτωση όμως είναι, η μη ανταπόκριση να μην είναι τυχαία αλλά συστηματική. Σε αυτή τη περίπτωση

---

<sup>10</sup> Κάποιοι θεωρούν τη μεροληψία επιλογής (*selection bias*) και την δειγματοληπτική μεροληψία (*sampling bias*) έννοιες ταυτόσημες. Κάποιοι άλλοι υποστηρίζουν ότι είναι διαφορετικές, με την δειγματοληπτική μεροληψία να οφείλεται σε σφάλματα στο αρχικό στάδιο συλλογής του δείγματος (π.χ η ικανότητα γενίκευσης του δείγματος στον πληθυσμό), ενώ η μεροληψία επιλογής αναφέρεται σε όλα τα υπόλοιπα λάθη που τυχόν προκύψουν στη συνέχεια. Γενικά το *sampling bias* και το *self-selection bias* θεωρούνται δύο υποκατηγορίες του *selection bias* (βλέπε Heckman, 1979).

<sup>11</sup> Περικομμένο λέγεται το δείγμα στο οποίο λείπουν όλες οι παρατηρήσεις μιας μεταβλητής, ενώ λογοκριμένο είναι το δείγμα στο οποίο λείπουν εν μέρει κάποιες παρατηρήσεις μίας μεταβλητής υπό κάποιες προϋποθέσεις (πχ οι ώρες εργασίας είναι γνωστές για τους εργαζόμενους αλλά όχι για τους άνεργους). Στα περικομμένα μοντέλα το πρόβλημα αφορά εξαρτημένες και ανεξάρτητες μεταβλητές, ενώ στα λογοκριμένα το πρόβλημα το έχει μόνο η εξαρτημένη μεταβλητή.

μπορεί να χρησιμοποιηθεί η μέθοδος του Heckman για τη ``συμπερασματική ρύθμιση`` (*imputation adjustment*)<sup>12</sup> των τιμών που λείπουν (Gilley and Leone, 1991).

Γενικά οι ερευνητές προσπαθούσαν να λύσουν το πρόβλημα των μη παρατηρούμενων τιμών στηριζόμενοι σε πολλές υποθέσεις και εστιάζοντας τη προσοχή τους κυρίως στα ερωτήματα: πόσες ώρες θα δούλευε κάποιος αν δεν ήταν άνεργος ή ποιος θα ήταν ο μισθός που θα έπαιρνε ένας άνεργος, δεδομένου ότι δούλευε. Όμως δεν έχει επιτευχθεί κάποια ομοφωνία. Ο Heckman (1993) κάνει λόγο για τη σοβαρότητα που πρέπει να δείχνει κανείς στο πρόβλημα των μη παρατηρούμενων τιμών και κατ' επέκταση και στη μεροληψία του δείγματος και αναφέρει ότι σε αυτό το θέμα δεν είναι εύκολο να υπάρξει μία ομοφωνία, καθώς διαφορετικές υποθέσεις μπορούν να είναι κάθε φορά κατάλληλες, ανάλογα με το πρόβλημα που αντιμετωπίζει ένας ερευνητής και τα διαφορετικά σύνολα δεδομένων που έχει στην κατοχή του. Επομένως το πρόβλημα του "συμπερασματικού μισθού" (*wage imputation problem*) καθώς και οποιασδήποτε άλλης λογοκριμένης μεταβλητής θα προκαλεί αντιδράσεις.

Τα περισσότερα δείγματα με λογοκριμένες εξαρτημένες μεταβλητές και ως επακόλουθο και η μεροληψία τέτοιων δειγμάτων, αντιμετωπίζονται με τη διόρθωση που πρότεινε ο Heckman (1976, 1979) και η οποία (μαζί με το μοντέλο tobit) αποτελεί την πιο διαδεδομένη μέθοδο μέχρι και σήμερα. Η μέθοδος του Heckman είναι γνωστή και ως μέθοδος δύο σταδίων/βημάτων, LIML (Limited Information Maximum Likelihood), μοντέλο Heckit, αλλά και Tobit τύπου 2 (type 2 Tobit model) ή γενικευμένο μοντέλο tobit (generalized tobit model). Ο όρος LIML προέκυψε ως αντιδιαστολή στον όρο FIML (Full Information Maximum Likelihood), καθώς όπως θα δούμε και παρακάτω, η μέθοδος του Heckman γίνεται κάτω από ορισμένες συνθήκες, ενώ η FIML που είναι η κλασική μέθοδος μεγίστης πιθανοφάνειας-ML δεν υπόκειται σε τέτοιους περιορισμούς. Ο λόγος που προέκυψε το μοντέλο Heckman ήταν ως μία εναλλακτική λύση στη μέθοδο FIML, λόγω της μεγάλης υπολογιστικής πολυπλοκότητας της τελευταίας, και του μεγάλου κόστους των οικονομετρικών πακέτων που απαιτούνταν εκείνη την εποχή για να εκτιμήσουν τη πλήρη συνάρτηση μεγίστης πιθανοφάνειας.

---

<sup>12</sup> Εκτός από τη συμπερασματική ρύθμιση, υπάρχουν και άλλες μέθοδοι που χρησιμοποιούνται στη δειγματοληψία για τη ρύθμιση της μη ανταπόκρισης (για περισσότερες πληροφορίες βλ. Χαρίσης και Κιόχος, 1997: κεφάλαιο 12)



## 4.2 Η μεθοδολογία των δύο σταδίων του Heckman – μοντέλο Heckit

Ο Heckman (1979) μελετά και παρομοιάζει το πρόβλημα της μεροληψίας του δείγματος ως ένα σφάλμα εξειδίκευσης που προέρχεται από την παράλειψη μιας σχετικής μεταβλητής. Υπάρχουν δύο παράγοντες που μπορούν να προκαλέσουν το πρόβλημα της μεροληψίας: Ο ένας είναι η αυτοεπιλογή (*self-selection*) που αφορά άτομα που από μόνα τους επιλέγουν να μουν σε μία ομάδα μελέτης και ο δεύτερος αφορά τη λανθασμένη επιλογή του δείγματος από τους ερευνητές. Η ανάλυση του Heckman (1979) αποδίδει μία συνεπή μέθοδο εκτίμησης η οποία εξαλείφει το σφάλμα εξειδίκευσης όταν έχουμε λογοκριμένα δείγματα και δίνει έτσι τη δυνατότητα στους ερευνητές να χρησιμοποιούν απλές μεθόδους παλινδρόμησης όπως αυτή των ελαχίστων τετραγώνων. Η διαδικασία του Heckman παρουσιάζεται εν συντομία με το παρακάτω παράδειγμα. Υποθέτουμε ότι έχουμε δύο εξαρτημένες μεταβλητές,  $Y_1$  (έστω ο μισθός ή οι ώρες εργασίας) και  $Y_2$  (ψευδομεταβλητή που δείχνει αν κάποιος εργάζεται ή όχι) όπου η  $Y_1$  είναι γνωστή μόνο όταν  $Y_2 > 0$ . Αντίθετα, όταν  $Y_2 \leq 0$  δεν έχουμε παρατηρήσεις για την  $Y_1$ . Άρα έχουμε ένα ολοκληρωμένο δείγμα μεγέθους  $N$  (για όλες τις παρατηρήσεις της  $Y_2$ ) και ένα υπο-δείγμα μεγέθους  $N_1$  (για τις παρατηρήσεις  $Y_2 > 0$ ). Φυσικά ισχύει ότι  $N_1 < N$ . Το μοντέλο αποτελείται από δύο εξισώσεις:

$$Y_{1i} = X_{1i}\beta_1 + U_{1i} \quad (1a)$$

$$Y_{2i} = X_{2i}\beta_2 + U_{2i} \quad (1b)$$

όπου  $X$  είναι ένα διάνυσμα ερμηνευτικών μεταβλητών το οποίο είναι γνωστό για όλες τις παρατηρήσεις της  $Y_2$  και  $U$  ο διαταρακτικός όρος. Αν δεν υπήρχε το πρόβλημα με τις ελλειπείς παρατηρήσεις τότε θα αναμέναμε ότι  $E(Y_{1i}|X_{1i}) = X_{1i}\beta_1$ . Όμως τώρα, δεδομένου ότι έχουμε ένα λογοκριμένο δείγμα, η αναμενόμενη συνάρτηση παλινδρόμησης θα είναι η εξής:

$$E(Y_{1i}|X_{1i}, Y_{2i} > 0) = X_{1i}\beta_1 + E(U_{1i}|Y_{2i} > 0) \Rightarrow$$
$$E(Y_{1i}|X_{1i}, Y_{2i} > 0) = X_{1i}\beta_1 + E(U_{1i}|U_{2i} > -X_{2i}\beta_2) \quad (2)$$

Εκτιμώντας το μοντέλο με τη μέθοδο ελαχίστων τετραγώνων (OLS) παραβιάζεται μία από τις βασικές υποθέσεις της OLS που λέει ότι το σφάλμα και οι ανεξάρτητες μεταβλητές δεν πρέπει να συσχετίζονται, αφού κοιτάζοντας τον δεύτερο όρο στο δεξί μέλος της εξίσωσης (2), βλέπουμε ότι ο διαταρακτικός όρος σχετίζεται με το διάνυσμα  $X_2$  των ανεξάρτητων μεταβλητών. Αυτή η συσχέτιση προκαλεί το πρόβλημα της μεροληψίας του δείγματος (Winship and Mare, 1992). Ο δεύτερος όρος στο δεξί μέλος της εξίσωσης (2), σύμφωνα με τον Heckman (1979), είναι ουσιαστικά η μεταβλητή που αν παραληφθεί θα έχουμε σφάλμα εξειδίκευσης με αποτέλεσμα οι OLS εκτιμητές του  $\beta_1$  να είναι μεροληπτικοί και ασυνεπείς.

Αν υποθέσουμε ότι τα σφάλματα  $U_{1i}$  και  $U_{2i}$  ακολουθούν μία κοινή κανονική κατανομή (*bivariate normal distribution*), τότε  $E(U_{1i}|U_{2i} > -X_{2i}\beta_2) = \frac{\sigma_{12}}{\sqrt{\sigma_{22}}} \lambda(Z_i)$  όπου  $\sigma_{12}$  η συνδιακύμανση των δύο σφαλμάτων,  $\sigma_{22}$  η διακύμανση του σφάλματος  $U_{2i}$  και  $\lambda(Z_i)$  το inverse Mill's ratio. Άρα η (2) γίνεται

$$Y_{1i} = X_{1i}\beta_1 + \frac{\sigma_{12}}{\sqrt{\sigma_{22}}} \lambda(Z_i) + \varepsilon_i \quad (3)$$

ή

$$Y_{1i} = X_{1i}\beta_1 + C \lambda(Z_i) + \varepsilon_i \quad \text{όπου } C = \frac{\sigma_{12}}{\sqrt{\sigma_{22}}}$$

Το inverse Mill's ratio είναι ο λόγος μεταξύ της συνάρτησης πυκνότητας πιθανότητας  $\varphi(Z)$  και της αθροιστικής συνάρτησης κατανομής  $\Phi(Z)$ . Δηλαδή  $\lambda(Z_i) = \frac{\varphi(Z_i)}{1-\Phi(Z_i)} = \frac{\varphi(Z_i)}{\Phi(-Z_i)}$  όπου

$$Z_i = -\frac{X_{2i}\beta_2}{\sqrt{\sigma_{22}}} \quad (4)$$

Το μοντέλο του Heckman για τη διόρθωση του σφάλματος/της μεροληψίας του δείγματος συνοψίζεται στα εξής βήματα: α) Με τη μέθοδο Probit και για ολόκληρο το δείγμα  $N$  εκτιμάμε τη πιθανότητα κάποιος να εργάζεται, δηλαδή εκτιμάμε την εξίσωση (1b). β) Από αυτή τη παλινδρόμηση παίρνουμε τον εκτιμημένο συντελεστή  $\widehat{\beta}_2$  και τον χρησιμοποιούμε για να εκτιμήσουμε τη μεταβλητή  $Z_i$  και ως εκ τούτου και τον συντελεστή  $\lambda_i$ , ο οποίος είναι το inverse Mill's ratio (τα περισσότερα προγράμματα εμφανίζουν άμεσα το inverse Mill's ratio). γ) Χρησιμοποιούμε το inverse Mill's ratio  $[\hat{\lambda}(Z_i)]$  ως ανεξάρτητη μεταβλητή στην εξίσωση (1a) και τρέχουμε την παλινδρόμηση με τη μέθοδο OLS. Με λίγα λόγια τρέχουμε τη παλινδρόμηση (3) για το υπο-δείγμα  $N_1$  στο οποίο όλες οι παρατηρήσεις της  $Y_1$  είναι γνωστές.

Η OLS εκτίμηση της εξίσωσης (3) δίνει συνεπείς και αμερόληπτους εκτιμητές των  $\beta_1$  και  $\sigma_{12}/\sqrt{\sigma_{22}}$  [δεδομένου ότι  $E(\varepsilon_i|Y_{2i}>0) = 0$ ], οι οποίοι είναι ασυμπτωτικά κανονικοί (Leung and Yu, 2000), αλλά είναι αναποτελεσματικοί γιατί υπάρχει πρόβλημα ετεροσκεδαστικότητας με τα σφάλματα  $\varepsilon_i$ , τα οποία ίσως και να συσχετίζονται. Γι' αυτό το λόγο θα πρέπει να χρησιμοποιήσουμε κάποιες μεθόδους για να πάρουμε τα σωστά τυπικά σφάλματα. Οι μέθοδοι που έχουν προταθεί είναι η GLS (Generalized Least Squares) (Heckman, 1976) και η μέθοδος του συνεπών σφαλμάτων του White (Lee, 1982). Η ετεροσκεδαστικότητα των σφαλμάτων  $\varepsilon_i$  οφείλεται στο γεγονός ότι σχετίζονται με τη μεταβλητή  $Z_i$  [και κατ' επέκταση με τη μεταβλητή  $X_{2i}$  σύμφωνα με τη σχέση (4)] όπως φαίνεται στο παρακάτω τύπο (Heckman, 1979)

$$E(\varepsilon_i^2 | X_{1i}, \lambda(Z_i), U_{2i} > -X_{2i}\beta_2) = \sigma_{11}[(1-\rho^2) + \rho^2[1+Z_i\lambda(Z_i) - \lambda^2(Z_i)]] \quad (5)$$

$$\text{όπου ο συντελεστής συσχέτισης } \rho = \frac{\sigma_{12}}{\sqrt{\sigma_{11}}\sqrt{\sigma_{22}}}, \quad (6)$$

$\sigma_{11}$  η διακύμανση του  $U_{1i}$  και

$$0 \leq 1+Z_i\lambda(Z_i) - \lambda^2(Z_i) \leq 1 \quad (7)$$

Η συνεπής εκτίμηση του  $\sigma_{11}$  γίνεται παίρνοντας από τη παλινδρόμηση (3) τα κατάλοιπα  $\hat{\varepsilon}_i$  και τον εκτιμημένο συντελεστή  $\hat{C} = \sigma_{12}/\sqrt{\sigma_{22}}$ , ο οποίος λόγω της σχέσης (6) ισοδυναμεί με

$$\hat{C} = \rho\sqrt{\sigma_{11}} \quad (8)$$

και χρησιμοποιώντας το τύπο:

$$\hat{\sigma}_{11} = \frac{\sum_{i=1}^{N^*} \hat{\varepsilon}_i^2}{N^*} - \frac{\hat{C}^2}{N^*} \sum_{i=1}^{N^*} \hat{\lambda}_i \hat{Z}_i - \hat{\lambda}_i^2, \text{ όπου } \lambda_i \equiv \lambda(Z_i)$$

Ο τύπος αυτός προκύπτει από τη διόρθωση που έκανε ο Greene (1981), βάζοντας τον όρο  $\hat{C}^2$  αντί του  $\hat{C}$  που είχε ο Heckman (1979). Ο εκτιμητής της πληθυσμιακής διακύμανσης  $\hat{\sigma}_{11}$  είναι συνεπής και έχει θετικό πρόσημο αφού ο όρος  $\hat{\lambda}_i \hat{Z}_i - \hat{\lambda}_i^2$  είναι αρνητικός [από τη σχέση (7) προκύπτει ότι  $-1 \leq \hat{\lambda}_i \hat{Z}_i - \hat{\lambda}_i^2 \leq 0$ ].

### 4.3 Πιθανά προβλήματα του μοντέλου Heckit

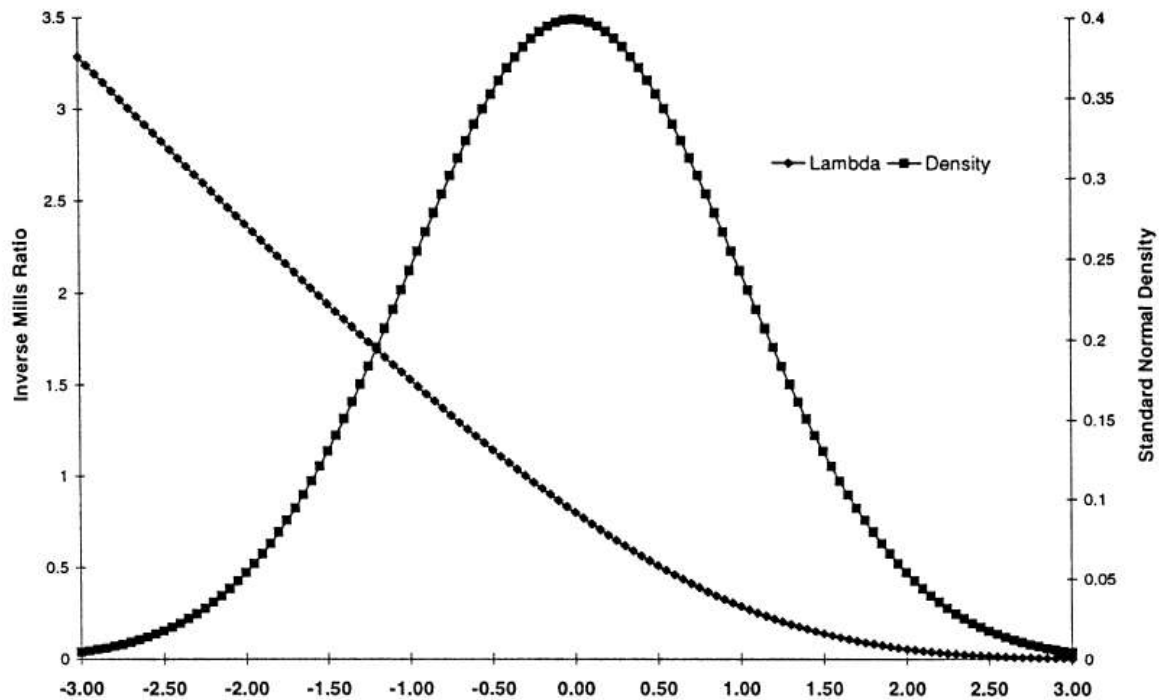
Το μοντέλο Heckit, όπως περιγράφηκε παραπάνω, αποτελεί σπουδαίο εύρημα στην αντιμετώπιση του sample selection bias και χρησιμοποιείται ευρέως στην εμπειρική έρευνα πολλών κλάδων της μικροοικονομίας με κυριότερους τους κλάδους της υγείας, της αγοράς εργασίας, αλλά και των επιχειρήσεων. Παρόλα αυτά, από τη στιγμή που πρωτοεμφανίστηκε στη βιβλιογραφία από τον Heckman (1976, 1979), δέχεται κριτικές από κάποιους ερευνητές λόγω των υποθέσεων του. Το μοντέλο στηρίζεται σε δύο βασικές υποθέσεις: τα σφάλματα  $U_{1i}$  και  $U_{2i}$  ακολουθούν μία κοινή κανονική κατανομή [ $h(U_{1i}|U_{2i}) \sim N(0,0, \sigma_{11}, \sigma_{22})$ ] και  $E(U_{1i}|U_{2i})$  είναι γραμμικό. Αν δεν ισχύουν αυτές οι υποθέσεις, το μοντέλο Heckit δεν βγάζει σωστά αποτελέσματα. Πολλοί ερευνητές θέλησαν να βρουν εναλλακτικές μεθόδους που να μην εξαρτώνται από τις υποθέσεις της κανονικότητας και γραμμικότητας. Για παράδειγμα ο Olsen (1980) έδειξε ότι η υπόθεση της κοινής κανονικότητας (*bivariate normality*) είναι επαρκής αλλά όχι αναγκαία. Δηλαδή η υπόθεση της κοινής κανονικότητας των δύο σφαλμάτων δεν χρειάζεται. Μόνο τα σφάλματα  $U_{2i}$  πρέπει να ακολουθούν την κανονική κατανομή και φυσικά πρέπει να ισχύει και η γραμμικότητα της υπό συνθήκη αναμενόμενης τιμής  $E(U_{1i}|U_{2i})$ . Στις περισσότερες εμπειρικές μελέτες μάλιστα, υποθέτουν ότι τα κατάλοιπα  $U_{2i}$  ακολουθούν την τυπική κανονική κατανομή  $N(0,1)$  και οποιαδήποτε υπόθεση περί μη

μηδενικό μέσο των καταλοίπων, αντιμετωπίζεται με έναν σταθερό όρο που σχεδόν πάντα υιοθετείται στο μοντέλο probit (Yamagata and Orme, 2005).

Επίσης, εκτός από τις αρχικές υποθέσεις του μοντέλου του Heckman, συζήτηση έχει γίνει αναφορικά και με τα πιθανά προβλήματα πολυσυγγραμμικότητας ανάμεσα στον συντελεστή  $\lambda(Z_i)$  (inverse Mill's ratio) και το διάνυσμα ερμηνευτικών μεταβλητών  $X_{1i}$  στην εξίσωση (3). Αν οι ανεξάρτητες μεταβλητές  $X_{1i}$  και  $Z_i$  (δηλαδή οι μεταβλητές  $X_{1i}$  και  $X_{2i}$ ) έχουν υψηλή συσχέτιση μεταξύ τους, τότε θα υπάρχει πρόβλημα πολυσυγγραμμικότητας μεταξύ του  $\lambda(Z_i)$  και του  $X_{1i}$  (Nawata, 1993; Nawata 1994; Nawata and Nagase, 1996). Πολλές φορές τα δύο διανύσματα  $X_{1i}$  και  $X_{2i}$  έχουν πολλές κοινές μεταβλητές, ενώ σε κάποιες περιπτώσεις είναι ταυτόσημα, αφού εμπεριέχουν ακριβώς τις ίδιες μεταβλητές.<sup>13</sup> Κατ' επέκταση η πολυσυγγραμμικότητα οφείλεται στο ότι το  $\lambda(Z_i)$  είναι μία σχεδόν γραμμική συνάρτηση της  $Z_i$  στο μεγαλύτερο εύρος τιμών της  $Z_i$  (Puhani, 2000). Στην Εικόνα 2 φαίνεται η μορφή που έχει η σχεδόν γραμμική συνάρτηση του  $\lambda(Z_i)$ . Ο Nawata (1993, 1994) υποστήριξε επίσης ότι στη παρουσία της πολυσυγγραμμικότητας, ο εκτιμητής του Heckman κρίνεται ακατάλληλος, ενώ αντίθετα ο FIML εκτιμητής είναι εύρωστος και πιο αξιόπιστος σε αυτή τη περίπτωση (Nawata, 1994; Nawata and Nagase, 1996). Οι Leung and Yu (1996, 2000) έκαναν κριτική στην άποψη του Nawata (1993, 1994), υποστηρίζοντας ότι ο εκτιμητής του Heckman ενδείκνυται ακόμα και όταν υπάρχει υψηλός βαθμός πολυσυγγραμμικότητας μεταξύ  $X_{1i}$  και  $X_{2i}$ , υπό την προϋπόθεση ότι η αναλογία των λογοκριμένων παρατηρήσεων είναι μικρή ή το εύρος τιμών του  $X_{2i}$  είναι μεγάλο. Αν δηλαδή ο αριθμός των παρατηρήσεων που λείπουν είναι μικρός σε σχέση με το μέγεθος ολόκληρου του δείγματος, ή υπάρχει μεγάλη μεταβλητότητα στις τιμές του  $X_{2i}$ , τότε ο εκτιμητής του Heckman (και κατ' επέκταση το μοντέλο Heckit) αποδίδει καλά ακόμα και όταν υπάρχει τέλεια πολυσυγγραμμικότητα ( $X_{1i}=X_{2i}$ ). Επομένως μία υψηλή συσχέτιση των  $X_{1i}$  και  $X_{2i}$  δεν συνεπάγεται απαραίτητα και πρόβλημα πολυσυγγραμμικότητας στην εξίσωση (3) μεταξύ  $\lambda(Z_i)$  και  $X_{1i}$ . Μεγάλη μεταβλητότητα στις τιμές  $X_{2i}$  σημαίνει μεγάλη διασπορά των εκτιμημένων τιμών  $Z_i$  στο δείγμα  $N_1$  (όπου  $Y_{2i}>0$ ). Με λίγα λόγια, σύμφωνα με τους Leung and Yu (2000), αν οι τιμές  $\hat{Z}_i$  (δηλαδή οι τιμές  $X_{2i}\hat{\beta}_2$  δεδομένου ότι  $\sigma_{22}=1$ ) και  $\lambda(\hat{Z}_i)$  είναι διασκορπισμένες σε όλο το μήκος της συνάρτησης  $\lambda(Z)$  ή μόνο στο δεξιό-μη γραμμικό κομμάτι (βλ. Εικόνα 2), τότε δεν θα υπάρχουν προβλήματα πολυσυγγραμμικότητας. Ο Nawata (1993) εξέτασε μόνο το γραμμικό κομμάτι της συνάρτησης και γι' αυτό το λόγο τα γενικά συμπεράσματά του περί πολυσυγγραμμικότητας και μη αξιόπιστου μοντέλου Heckit δεν είναι εντελώς σωστά.

<sup>13</sup> Στην ειδική περίπτωση όπου  $X_{1i} \equiv X_{2i}$  προκύπτει το μοντέλο Tobit.

Εικόνα 2: Η οιονεί γραμμικότητα του inverse mills λόγου



Πηγή: Puhani (2000)

Η ταυτοποίηση του μοντέλου Heckit βασίζεται στην εύρεση κάποιων ερμηνευτικών μεταβλητών που εισέρχονται στην εξίσωση probit (εξίσωση 1b) αλλά δεν συμπεριλαμβάνονται στην εξίσωση παλινδρόμησης (εξίσωση (3)). Με λίγα λόγια θα πρέπει να βρούμε κάποιες ανεξάρτητες μεταβλητές που επηρεάζουν την απόφαση κάποιου να εισέλθει στην αγορά εργασίας, αλλά δεν επηρεάζουν το πόσες ώρες θα εργαστεί. Στην πράξη είναι λίγο δύσκολο να βρεθούν αυτές οι μεταβλητές και γι' αυτό μερικές φορές το μοντέλο Heckit εκτιμάται με το ίδιο σύνολο μεταβλητών και στις δύο εξισώσεις (δηλαδή  $X_{1i}=X_{2i}$ ). Σε αυτή τη περίπτωση, η ταυτοποίηση του μοντέλου Heckit βασίζεται στη μη γραμμικότητα του inverse Mill's ratio, αφού το inverse Mill's ratio είναι συνάρτηση του γραμμικού όρου  $X_{2i}\beta_2$  από την εξίσωση probit (βλ. Jones, 2007: chapter 7).

Συνοπτικά (βλ. Jones, 2007: chapter 7), υψηλή πολυσυγγραμμικότητα είναι πιθανόν να υπάρχει όταν α) οι ερμηνευτικές μεταβλητές που αποκλείονται από την εξίσωση παλινδρόμησης (3) είναι λίγες ή καμία β) ο βαθμός της μη ανταπόκρισης είναι μεγάλος (δηλαδή η αναλογία των μη παρατηρούμενων τιμών σε σχέση με τις παρατηρούμενες τιμές είναι μεγάλη) γ) η μεταβλητότητα των ερμηνευτικών μεταβλητών  $X_{2i}$  στην εξίσωση probit (1b) είναι χαμηλή δ) ο βαθμός των μεταβολών που μπορούμε να ερμηνεύσουμε στην εξίσωση probit είναι μικρός.

Ο Paarsch (1984) μέσω πειραμάτων Monte Carlo, θέλησε να συγκρίνει διάφορα λογοκριμένα μοντέλα παλινδρομήσεων (*censored regression models*) κάτω από διαφορετικές συνθήκες. Οι εκτιμητές που συγκρίνει είναι ο ML εκτιμητής του μοντέλου tobit, ο εκτιμητής δύο σταδίων του Heckman, ο ols εκτιμητής χρησιμοποιώντας μόνο το δείγμα με τις θετικές παρατηρήσεις, και ο LAD (least absolute deviation) εκτιμητής που πρότεινε ο Powell (1984) για λογοκριμένα δείγματα. Οι εναλλακτικές συνθήκες αφορούν: διαφορετικές κατανομές (Laplace, Cauchy, normal distribution), διαφορετικά μεγέθη δείγματος (N=50, 100, 200) και διαφορετικό βαθμό λογοκρισίας-μη ανταπόκρισης (25%, 50%). Παρακάτω έχουμε ενδεικτικά τον πίνακα αποτελεσμάτων που αφορά τη κανονική κατανομή, μέγεθος δείγματος 200 και λογοκρισία 50%, και τον αντίστοιχο πίνακα με λογοκρισία 25%. Ο συντελεστής  $a$  είναι ο σταθερός όρος και ο  $b$  ο συντελεστής του διανύσματος των ανεξάρτητων μεταβλητών. Οι στήλες αντιστοιχούν στα εξής στατιστικά μέτρα:

Mean=η μέση τιμή της αθροιστικής συνάρτησης κατανομής

S.d=η τυπική απόκλιση της αθροιστικής συνάρτησης κατανομής

L.q=το κάτω τεταρτημόριο της αθροιστικής συνάρτησης κατανομής

Median=ο διάμεσος της αθροιστικής συνάρτησης κατανομής

U.q=το πάνω τεταρτημόριο της αθροιστικής συνάρτησης κατανομής

Στις παρενθέσεις είναι οι τυπικές αποκλίσεις κάθε στατιστικού μέτρου που βασίζονται σε εκτιμήσεις bootstrap.

**Πίνακας 1: Αποτελέσματα των 4 εκτιμήσεων**

Normal – sample size = 200, fifty percent censoring.							
		Truth	Mean	S.d.	L.q.	Median	U.q.
Tobit	$a$	-10.00	-10.26 (0.20)	2.03 (0.14)	-11.87 (0.50)	-10.32 (0.27)	-8.95 (0.24)
	$b$	1.00	1.02 (0.02)	0.15 (0.01)	0.91 (0.02)	1.00 (0.02)	1.11 (0.03)
Powell	$a$	-10.00	-12.48 (0.76)	7.58 (0.63)	-15.55 (1.67)	-10.72 (0.74)	-7.04 (0.55)
	$b$	1.00	1.15 (0.05)	0.47 (0.04)	0.78 (0.05)	1.04 (0.04)	1.37 (0.11)
Heckman	$a$	-10.00	-12.31 (3.12)	31.24 (3.36)	-29.53 (3.63)	-14.06 (3.38)	5.46 (2.87)
	$b$	1.00	1.08 (0.13)	1.27 (0.13)	0.27 (0.13)	1.21 (0.15)	1.80 (0.13)
POLS	$a$	-10.00	4.23 (0.14)	1.38 (0.11)	3.40 (0.20)	4.07 (0.15)	5.19 (0.24)
	$b$	1.00	0.41 (0.01)	0.14 (0.01)	0.33 (0.02)	0.43 (0.02)	0.49 (0.02)

Normal – sample size = 200, twenty-five percent censoring.

		Truth	Mean	S.d.	L.q.	Median	U.q.
Tobit	<i>a</i>	–2.09	–2.15 (0.17)	1.66 (0.12)	–3.60 (0.27)	–1.95 (0.16)	–1.01 (0.17)
	<i>b</i>	1.00	1.01 (0.01)	0.13 (0.01)	0.91 (0.01)	1.00 (0.02)	1.11 (0.03)
Powell	<i>a</i>	–2.09	–2.59 (0.25)	2.54 (0.20)	–3.52 (0.43)	–2.08 (0.25)	–1.05 (0.71)
	<i>b</i>	1.00	1.04 (0.02)	0.19 (0.01)	0.91 (0.02)	1.00 (0.01)	1.15 (0.02)
Heckman	<i>a</i>	–2.09	–3.00 (1.44)	14.41 (1.19)	–12.59 (2.01)	–1.19 (1.77)	7.12 (1.70)
	<i>b</i>	1.00	1.02 (0.09)	0.90 (0.05)	0.56 (0.10)	0.96 (0.09)	1.50 (0.13)
POLS	<i>a</i>	–2.09	6.05 (0.12)	1.18 (0.08)	5.29 (0.17)	6.15 (0.14)	6.81 (0.13)
	<i>b</i>	1.00	0.60 (0.01)	0.10 (0.01)	0.52 (0.02)	0.60 (0.01)	0.67 (0.02)

Πηγή: Paarsch (1984)

Κοιτώντας τις δύο πρώτες στήλες βλέπουμε ότι η μέση τιμή του εκτιμητή *ols* απέχει πολύ από τη πραγματική, που σημαίνει ότι η μέθοδος *ols* δεν είναι κατάλληλη. Αντίθετα οι άλλοι τρεις εκτιμητές είναι πιο κοντά στη πραγματική τιμή, με αυτόν του *tobit* να φαίνεται πιο ακριβής από όλους. Όσον αφορά τα μεγάλα δείγματα (πάνω από 100 παρατηρήσεις), ο εκτιμητής που πρότεινε ο Powell (1984) για λογοκριμένες εξαρτημένες μεταβλητές είναι συνεπής για μία πληθώρα κατανομών καταλοίπων και είναι επίσης εύρωστος στην ετεροσκεδαστικότητα. Αυτή η ευρωστία τον κάνει ισχυρότερο και από την εκτίμηση *tobit*, η οποία δίνει μη συνεπείς εκτιμητές όταν τα σφάλματα είναι ετεροσκεδαστικά.

#### 4.4 Διαγνωστικοί έλεγχοι υποθέσεων για την ύπαρξη της μεροληψίας επιλογής

Ο Heckman (1979) προτείνει να ελέγχεται η ύπαρξη της μεροληψίας επιλογής μέσω του *t*-τεστ για τον συντελεστή του  $\lambda$ , υπό της μηδενικής υπόθεσης για μη ύπαρξη μεροληψίας ( $H_0$ : no selection bias). Δηλαδή ο έλεγχος στατιστικής σημαντικότητας είναι ο εξής:

$H_0$ :  $C=0$

$H_1$ :  $C \neq 0$

όπου  $C$  ο συντελεστής του  $\lambda$  από την εξίσωση (3). Αν ο συντελεστής του  $\lambda$  είναι στατιστικά σημαντικός, τότε υπάρχει πρόβλημα μεροληψίας. Ο Melino (1982) έδειξε ότι ο έλεγχος με το *t*-statistic ισοδυναμεί με τον έλεγχο LM (Lagrange Multiplier). Συγκεκριμένα ο

πολλαπλασιαστικής Lagrange ισούται με το τετράγωνο του t-statistic του συντελεστή C ( $LM=t_c^2$ ) και λόγω των ασυμπτωτικών του ιδιοτήτων, είναι αποτελεσματικός για μεγάλα δείγματα και μπορεί να συγκριθεί με τα τεστ Wald και LR (Likelihood Ratio). Τα τεστ ελέγχου t του Heckman (1979) και LM του Melino (1982), γενικά είναι εύρωστα (*robust*) στη μη κανονικότητα και μπορούν να χρησιμοποιηθούν όταν το πρόβλημα πολυσυγγραμμικότητας δεν είναι πολύ μεγάλο (Yamagata and Orme, 2005; Leung and Yu, 1996). Αντίθετα, οι Nawata and McAleer (2001) υπό την μηδενική υπόθεση της μη ύπαρξης μεροληψίας και υποθέτοντας κανονικότητα, σύγκριναν τα τεστ LR, LM και Wald και διαπίστωσαν ότι όταν υπάρχει σοβαρό πρόβλημα πολυσυγγραμμικότητας, το LR τεστ παραμένει εύρωστο, ενώ τα άλλα δύο δεν είναι τόσο αξιόπιστα. Οι Yamagata and Orme (2005) τρέχοντας εναλλακτικά πειράματα με προσομοιώσεις Monte Carlo, κατέληξαν στο ότι το t-test (Heckman, 1979) και το LM-test (Melino, 1982) είναι προτιμότερο να χρησιμοποιούνται όταν δεν υπάρχει πολυσυγγραμμικότητα, ενώ όταν υπάρχει, είναι προτιμότερο να επιλέγεται το τεστ LR μέσω μιας εκτίμησης μεγίστης πιθανοφάνειας-ML. Η άποψη αυτή περί καταλληλότερου τεστ ελέγχου συμβαδίζει με την άποψη των Nawata and Nagase (1996) περί καταλληλότερης μεθόδου εκτίμησης, οι οποίοι συγκρίνοντας τον ML εκτιμητή  $\beta_1$  (βλ. εξίσωση 3) με τον αντίστοιχο OLS του Heckman, συμπέραναν ότι ο εκτιμητής μεγίστης πιθανοφάνειας υπερτερεί του εκτιμητή του Heckman κυρίως όταν υπάρχει πολυσυγγραμμικότητα. Όμως σε περίπτωση μη κανονικότητας πρέπει να είμαστε επιφυλακτικοί καθώς η εκτίμηση μεγίστης πιθανοφάνειας μπορεί να ερμηνεύσει τη μη κανονικότητα ως sample selection bias (Yamagata and Orme, 2005). Γενικά για τη χρήση του τεστ LR (και κατ' επέκταση της μεθόδου ML) απαιτείται η υπόθεση της κοινής κανονικότητας (*bivariate normality*), ενώ τα t-test και LM-test είναι ασυμπτωτικά έγκυρα κάτω από λιγότερο αυστηρές υποθέσεις (όπως πχ η απλή κανονικότητα).

Ένας ακόμη έλεγχος για sample selection bias είναι και ο συντελεστής προσδιορισμού  $R^2$  που παίρνουμε από μία βοηθητική παλινδρόμηση του εκτιμημένου inverse Mill's ratio πάνω σε όλες τις μεταβλητές του διανύσματος  $X_{1i}$  (Nelson, 1984; Leung and Yu, 1996; Yamagata and Orme, 2005). Όσο μεγαλύτερος είναι ο συντελεστής προσδιορισμού τόσο μεγαλύτερο είναι το πρόβλημα της πολυσυγγραμμικότητας. Το ίδιο ισχύει και για τον προσαρμοσμένο συντελεστή προσδιορισμού. Ως επακόλουθο, μπορούμε να χρησιμοποιήσουμε και τον συντελεστή διόγκωσης της διακύμανσης (*variance inflation factor*) (Leung and Yu, 1996). Ο συντελεστής αυτός αποτελεί ένα μέτρο εντοπισμού της πολυσυγγραμμικότητας και δίνεται από τον τύπο  $VIF = 1/(1-R^2)$ . Όσο μεγαλύτερη είναι η τιμή του VIF τόσο μεγαλύτερο και το πρόβλημα της πολυσυγγραμμικότητας (βλ. Χάλκος,



2011: Κεφ. 6). Δεν υπάρχει κάποια συγκεκριμένη κριτική τιμή, αλλά γενικά αν  $VIF > 10$  τότε έχουμε πρόβλημα. Οι Leung and Yu (1996) προτείνουν και έναν καλύτερο έλεγχο από το συντελεστή προσδιορισμού, που βασίζεται στις ιδιοτιμές της μήτρας  $X'X$  της εξίσωσης (3). Συγκεκριμένα ο έλεγχος αυτός αφορά έναν αριθμό, ο οποίος ονομάζεται condition number, και προκύπτει διαιρώντας τη μεγαλύτερη ιδιοτιμή με τη μικρότερη και παίρνοντας τη τετραγωνική ρίζα αυτού του κλάσματος ( $\text{condition number} = \sqrt{\max \text{eigenvalue} / \min \text{eigenvalue}}$ ). Όσο μεγαλύτερο είναι το condition number τόσο πιθανότερο είναι να υπάρχει πολυσυγγραμμικότητα. Δεν υπάρχει κάποια συγκεκριμένη κριτική τιμή, αλλά εμπειρικά με προσομοιώσεις Monte Carlo φαίνεται ότι μία τιμή άνω της οποίας παρουσιάζεται το πρόβλημα, είναι η 30.

Ο Olsen (1982) πρότεινε ένα τεστ ελέγχου που βασίζεται στη κατανομή των καταλοίπων. Συγκεκριμένα παρουσιάζει δύο παραλλαγές του τεστ (υπό τη μηδενική υπόθεση  $H_0$ : no selection bias) οι οποίες βασίζονται στη σύγκριση των καταλοίπων από δύο ομάδες ατόμων που προέρχονται από το δείγμα  $N_1$ . Για παράδειγμα από ένα δείγμα  $N=7228$ , με άντρες και γυναίκες ηλικίας 13 έως 20 χρονών, μόνο 2733 εργάζονταν και δήλωσαν το μισθό τους. Δηλαδή  $N_1=2733$ . Το δείγμα  $N_1$  το χωρίζει σε δύο ομάδες, με τη μία ομάδα να περιλαμβάνει τους 954 με τις μεγαλύτερες προβλεπόμενες πιθανότητες να εργάζονται, και την άλλη ομάδα τους 954 με τις χαμηλότερες προβλεπόμενες πιθανότητες να εργάζονται. Ο λόγος που αφαιρεί τις μεσαίες 825 παρατηρήσεις είναι για να δυναμώσει το τεστ ελέγχου<sup>14</sup>. Οι προβλεπόμενες πιθανότητες να εργάζεται κάποιος λαμβάνονται από την εκτίμηση της εξίσωσης (1b) είτε με το μοντέλο Probit είτε με το LPM (linear probability model). Στη συνέχεια εκτιμάει με OLS την εξίσωση  $Y_{1i} = X_i\beta_1 + U_{1i}$  για κάθε μία ομάδα, όπου το διάνυσμα  $X_i$  περιλαμβάνει όλα τα στοιχεία των διανυσμάτων  $X_{1i}$  και  $X_{2i}$ . Το τεστ συγκρίνει τα κατάλοιπα από τις παλινδρομήσεις των δύο ομάδων. Για να ισχύει η μηδενική υπόθεση, θα πρέπει η κατανομή των καταλοίπων να είναι ίδια. Αντιθέτως, αν τα κατάλοιπα των δύο ομάδων έχουν διαφορετική κατανομή, τότε υπάρχει selection bias.

Τέλος παρουσιάζουμε ένα εμπειρικό παράδειγμα των Yamagata and Orme (2005) με εναλλακτικά τεστ ελέγχου. Η μηδενική υπόθεση εδώ είναι μία παραλλαγή της μηδενικής υπόθεσης του Heckman (1979). Αντί για  $H_0: C=0$ , ελέγχουν για  $H_0: \rho=0$ , όπου  $\rho$  ο συντελεστής συσχέτισης μεταξύ  $U_{1i}$  και  $U_{2i}$  (βλ. σχέση (6)). Η υπόθεση αυτή στηρίζεται στην εξίσωση παλινδρόμησης  $Y_{1i} = X_{1i}\beta_1 + \rho\sqrt{\sigma_{11}} * \lambda(Z_i) + \varepsilon_i$  (9), η οποία προκύπτει από αντικατάσταση της σχέσης (6) στην εξίσωση παλινδρόμησης (3). Ουσιαστικά οι δύο υποθέσεις είναι

<sup>14</sup> Η προσέγγιση που ακολουθεί ο Olsen (1982) μοιάζει με το τεστ ελέγχου Goldfeld-Quandt για ετεροσκεδαστικότητα.

εννοιολογικά ίδιες, αφού από τη σχέση (8) έχουμε ότι  $C=0 \Rightarrow \rho\sqrt{\sigma_{11}}=0 \Rightarrow \rho=0$  (αφού  $\sigma_{11} \neq 0$ ). Άρα και ο Heckman με τη μηδενική στατιστική σημαντικότητα του C, υπονοεί ουσιαστικά τη μηδενική στατιστική σημαντικότητα του συντελεστή συσχέτισης  $\rho$ . Οι Yamagata and Orme (2005) χρησιμοποίησαν τα δεδομένα του Mroz (1987) που περιλαμβάνουν 753 λευκές παντρεμένες γυναίκες, εκ των οποίων οι 428 εργάζονται. Το μοντέλο στηρίζεται στις αρχικές εξισώσεις (1a) και (1b), με  $Y_{2i}=1$  αν η γυναίκα εργάζεται και 0 σε αντίθετη περίπτωση, και  $Y_{1i}$  είναι ο φυσικός λογάριθμος του μισθού (*wage rate*) των γυναικών που εργάζονται. Οι συγγραφείς επιπλέον χρησιμοποιούν το ίδιο σύνολο ανεξάρτητων μεταβλητών στις δύο εξισώσεις (δηλαδή  $X_{1i}=X_{2i}$ ). Συνολικά χρησιμοποιούν 17 ανεξάρτητες μεταβλητές συμπεριλαμβανομένου του σταθερού όρου. Ενδεικτικά κάποιες από αυτές είναι: αριθμός παιδιών κάτω από 6 χρονών, αριθμός παιδιών μεταξύ 6 και 18 χρονών, οικογενειακό εισόδημα, ποσοστό ανεργίας της χώρας διαμονής, ηλικία, έτη εκπαίδευσης, τα τετράγωνα και οι κύβοι των δύο τελευταίων. Παρακάτω, στον πίνακα 1, φαίνονται τα αποτελέσματα πέντε εναλλακτικών τεστ ελέγχου. Τα τεστ  $t_{HG}^2$  και  $t_1^2$  είναι δύο παραλλαγές του t-test του Heckman, με το καθένα να βασίζεται σε διαφορετικά τυπικά σφάλματα του εκτιμητή  $\widehat{\rho\sqrt{\sigma_{11}}}$  από τη παλινδρόμηση (9). Το υπόλοιπα τρία τεστ είναι τα γνωστά Lagrange Multiplier, Likelihood ratio και Wald, που παίρνουμε από τη μέθοδο μεγίστης πιθανοφάνειας. Επίσης ο προσαρμοσμένος συντελεστής προσδιορισμού από τη παλινδρόμηση του εκτιμημένου *inverse Mill's ratio* (από την εξίσωση (9)) πάνω σε όλες τις ανεξάρτητες μεταβλητές είναι 0,98 που σημαίνει ότι υπάρχει μεγάλο πρόβλημα πολυσυγγραμμικότητας.

**Πίνακας 2: Έλεγχοι για sample selection bias**

	Statistic	p-value
$t_{HG}^2$	0.649	[0.4205]
$t_1^2$	0.652	[0.4195]
$LM_{AE}$	0.679	[0.4099]
$LR$	8.981	[0.0027]
$W$	219.029	[0.0000]

Πηγή: Yamagata and Orme (2005)

Στη πρώτη στήλη έχουμε τη τιμή των τεστ και στη δεύτερη τις τιμές P. Από τις p-values βλέπουμε ότι τα τρία πρώτα τεστ δέχονται τη μηδενική υπόθεση ( $H_0: \rho=0$ ) που σημαίνει ότι

δεν υπάρχει πρόβλημα μεροληψίας του δείγματος. Αντίθετα, σύμφωνα με τα τεστ LR και W πρέπει να απορρίψουμε τη μηδενική υπόθεση. Λόγω του μεγάλου προβλήματος πολυσυγγραμμικότητας που υπάρχει, τα δύο πρώτα τεστ που βασίζονται στη μεθοδολογία του Heckman δεν είναι αξιόπιστα. Επίσης είδαμε προηγουμένως ότι ούτε τα τεστ LM και W είναι αξιόπιστα όταν υπάρχει πολυσυγγραμμικότητα (Nawata and McAleer, 2001). Το μόνο εύρωστο τεστ είναι το LR, που στο συγκεκριμένο παράδειγμα απορρίπτει τη μηδενική υπόθεση ακόμα και σε επίπεδο στατιστική σημαντικότητας 1%. Άρα οι Yamagata and Orme (2005) καταλήγουν στο συμπέρασμα ότι μάλλον υπάρχει πρόβλημα μεροληψίας.

## Κεφάλαιο 5

### Εμπειρική εκτίμηση της συνάρτησης προσφοράς εργασίας

Στο κεφάλαιο αυτό γίνεται μία προσπάθεια να υποδειγματοποιηθεί εμπειρικά ένα μοντέλο προσφοράς εργασίας και να εκτιμηθούν οι αντίστοιχες ελαστικότητες προσφοράς εργασίας. Στο πρώτο μέρος παρουσιάζονται τα δεδομένα της έρευνας και στο δεύτερο μέρος γίνεται μία περιγραφή των βασικών μεταβλητών που θα χρησιμοποιηθούν στην εμπειρική ανάλυση. Στο τρίτο μέρος παρουσιάζονται τα ευρήματα της μελέτης.

#### 5.1 Δεδομένα

Τα δεδομένα αντλήθηκαν από την έρευνα που πραγματοποιήθηκε τη περίοδο του 2004 από τη Metron Analysis σε συνεργασία με το Τμήμα Μεθοδολογίας, Ιστορίας και Θεωρίας της Επιστήμης του Πανεπιστημίου Αθηνών, στα πλαίσια του ευρωπαϊκού προγράμματος SOCIOLD. Οι περιοχές μελέτης ήταν η Αθήνα, ο Πειραιάς, η Θεσσαλονίκη και μεγάλα αστικά κέντρα (Πάτρα, Ηράκλειο, Λάρισα, Βόλος κ.α.). Ο πληθυσμός έρευνας ήταν άτομα ηλικίας 45-65 ετών που μιλούν επαρκώς την ελληνική γλώσσα, ανεξαρτήτως εθνικότητας και φυλής, και είναι κάτοικοι της περιοχής μελέτης. Οι ερωτήσεις και κατ' επέκταση οι μεταβλητές είναι πολλές και ποικίλες και πιάνουν ένα ευρύ φάσμα δημογραφικών, κοινωνικών, προσωπικών και εργασιακών χαρακτηριστικών. Οι παρατηρήσεις είναι συνολικά 1001, από τις οποίες οι 438 είναι άντρες και οι 563 γυναίκες. Το δείγμα αποτελείται από μισθωτούς πλήρους και μερικής απασχόλησης, αυτοαπασχολούμενους, άτομα που ψάχνουν για εργασία (άνεργοι), άτομα που δεν ψάχνουν για εργασία, συνταξιούχους, άτομα που ασχολούνται με τα οικιακά και άτομα ανίκανα για εργασία. Στον παρακάτω πίνακα φαίνεται λεπτομερώς η κατανομή των παρατηρήσεων σε καθεμία ομάδα του δείγματος.

**Πίνακας 3: Ομαδοποιημένες παρατηρήσεις του δείγματος**

<b>Ομάδες δείγματος</b>	<b>Αριθμός παρατηρήσεων</b>
<b>Μισθωτοί με πλήρη απασχόληση</b>	265
<b>Μισθωτοί με μερική απασχόληση</b>	39
<b>Αυτοαπασχολούμενοι</b>	114
<b>Άνεργοι</b>	46
<b>Ηθελημένα άνεργοι</b>	22
<b>Συνταξιούχοι</b>	266
<b>Νοικοκυρές</b>	239
<b>Ανάπηροι</b>	10
<b>ΣΥΝΟΛΟ</b>	1001

Από τις παραπάνω ομάδες εξαιρούνται από την ανάλυση οι τέσσερις τελευταίες, δηλαδή τα άτομα που δεν ψάχνουν για εργασία, οι συνταξιούχοι, οι νοικοκυρές και οι ανάπηροι. Ο λόγος της εξαίρεσης αυτής είναι ότι το αντικείμενο μελέτης είναι η αγορά εργασίας και οι τέσσερις αυτές ομάδες δεν την επηρεάζουν, καθώς δεν ανήκουν σε αυτήν (Bishop et al., 2009; MaCurdy et al., 1990; Mroz, 1987). Επομένως το αρχικό μας δείγμα αποτελείται από 464 παρατηρήσεις (418 εργαζόμενοι + 46 άνεργοι). Πολλοί ερευνητές εξαιρούν επίσης και τους άνεργους και συμπεριλαμβάνουν στο δείγμα τους μόνο τα άτομα που εργάζονται. Στη περίπτωση όμως αυτή μπορεί να προκύψει ένα μεροληπτικό δείγμα, όπως είδαμε στο κεφάλαιο 4. Θα μπορούσαμε και εμείς εξ' αρχής να μην τους συμπεριλάβουμε στη μελέτη μας γιατί είναι λίγες οι παρατηρήσεις σε σχέση με ολόκληρο το δείγμα (46 στις 464 παρατηρήσεις) και πιθανόν να μην επηρεάζουν την αξιοπιστία των τελικών αποτελεσμάτων. Παρόλα αυτά όμως συμπεριλαμβάνονται και με τη βοήθεια του μοντέλου Heckit θα εξεταστεί κατά πόσο προκαλούν προβλήματα μεροληψίας.

Επίσης, οι περισσότεροι συγγραφείς τείνουν να αφήνουν εκτός ανάλυσης και τους αυτοαπασχολούμενους (Donni and Moreau, 2007). Λίγες είναι οι έρευνες που εξετάζουν μεταξύ των άλλων αυτά τα μέλη του εργατικού δυναμικού (Bishop et al., 2009; Kuroda and Yamamoto, 2008; Co et al., 2005; Parker et al., 2005; Schaffner, 2002; Maglad, 1998; Sharif, 1991; Nam, 1991), είτε μελετώντας μόνο κάποια τμήματα της αγοράς εργασίας, όπως εξισώσεις ωρών εργασίας (intensive margin) ή τη πιθανότητα συμμετοχής στο εργατικό δυναμικό (extensive margin), είτε γενικότερα μοντέλα προσφοράς εργασίας. Από αυτές τις μελέτες οι Parker et al. (2005) μελετάνε αποκλειστικά και μόνο τη συμπεριφορά των

αυτοαπασχολούμενων, ενώ οι Co et al. (2005) μελετάνε ταυτόχρονα τη πιθανότητα κάποιος να συμμετέχει στην αγορά εργασίας ή όχι (*participation selection*) και τη πιθανότητα κάποιος να επιλέξει την αυτοαπασχόληση σε σχέση με τη μισθωτή εργασία (*self-employment selection*) μέσω ενός μοντέλου Heckit στο οποίο το πρώτο βήμα είναι ένα bivariate probit model που δίνει δύο inverse mill's ratio, τα οποία υπεισέρχονται στο δεύτερο βήμα της ols παλινδρόμησης. Μία παρόμοια τακτική ακολουθεί και η El-Hamidi (2003). Ο λόγος που κάποιοι αφήνουν εκτός ανάλυσης τους αυτοαπασχολούμενους είναι ότι τις περισσότερες φορές δεν υπάρχουν διαθέσιμα δεδομένα όσον αφορά τις ώρες εργασίας ή το μισθό τους, αλλά ακόμα και να υπάρχουν, ίσως να μην αντιπροσωπεύουν τις πραγματικές ώρες-μισθό ή να έχουν μετρηθεί λανθασμένα.

Στην ερώτηση “πόσες περίπου ώρες τη μέρα/εβδομάδα/μήνα εργαστήκατε αυτό το έτος”, είναι δύσκολο να απαντήσει ένας αυτοαπασχολούμενος γιατί είναι πιθανόν να μην θυμάται. Από την άλλη μεριά, αν οι αυτοαπασχολούμενοι αποτελούν ένα μεγάλο μέρος του δείγματος και ταυτόχρονα υπάρχουν και όλα τα απαραίτητα δεδομένα για αυτούς, τότε δεν υπάρχει λόγος να αφήνονται εκτός ανάλυσης (Bishop et al., 2009). Στην συγκεκριμένη έρευνα της Metron Analysis έγιναν ξεχωριστά ερωτήσεις στους αυτοαπασχολούμενους όσον αφορά τις ώρες εργασίας τους και τα ετήσια καθαρά έσοδα από την εν λόγω επιχείρησή τους. Για αυτό το λόγο συμπεριλαμβάνονται στη μελέτη μας, αφού είναι και αρκετοί σε αριθμό και αν μείνουν εκτός ανάλυσης θα μειωθεί αρκετά το μέγεθος του δείγματος.

## 5.2 Ανάλυση και περιγραφή των μεταβλητών

### 5.2.1 Κύριες μεταβλητές

Οι δύο κύριες μεταβλητές που είναι απαραίτητες στην εξαγωγή των ελαστικοτήτων είναι οι ώρες εργασίας και ο μισθός. Στην έρευνα της Metron Analysis οι ερωτήσεις αναφορικά με τις δύο αυτές μεταβλητές γίνονται ξεχωριστά για τους μισθωτούς (πλήρους και μερικής απασχόλησης) και τους αυτοαπασχολούμενους. Όλοι οι ερωτώμενοι που είναι μισθωτοί απαντούν σε εβδομαδιαία βάση αναφορικά με το πόσες ώρες εργάζονται (συμπεριλαμβανομένου και των υπερωριών). Αντίθετα, αναφορικά με το μισθό, οι ερωτώμενοι μισθωτοί έχουν τρεις εναλλακτικές δυνατότητες απάντησης. Μπορούν να απαντήσουν σε εβδομαδιαία, δεκαπενθήμερη και μηνιαία βάση, ανάλογα δηλαδή με το κάθε πότε πληρώνονται. Στη συνέχεια γίνεται η ερώτηση: πόσες είναι οι καθαρές αποδοχές του από την εργασία του (δηλαδή μετά την αφαίρεση των φορολογικών και ασφαλιστικών εισφορών). Με λίγα λόγια, ο (καθαρός) μισθός που έχει δηλώσει ο κάθε μισθωτός μπορεί να

είναι εβδομαδιαίος, δεκαπενθήμερος ή μηνιαίος. Η συντριπτική πλειοψηφία απάντησαν σε δεκαπενθήμερη και μηνιαία βάση και γι' αυτό το λόγο επιλέχθηκε μία εκ των δύο ως βάση αναφοράς. Συγκεκριμένα η χρονική βάση αναφοράς είναι ο μήνας. Επομένως οι εβδομαδιαίοι και οι δεκαπενθήμεροι μισθοί θα πρέπει να μετατραπούν σε μηνιαίους. Οι σχετικοί υπολογισμοί που έλαβαν χώρα προκειμένου να επιτευχθεί αυτό, ήταν οι εξής: οι δεκαπενθήμεροι μισθοί πολλαπλασιάστηκαν με τον αριθμό 2 και οι εβδομαδιαίοι μισθοί πολλαπλασιάστηκαν με τον αριθμό 4,3. Ο λόγος που πολλαπλασίασα με 4,3 και όχι με 4 είναι για να υπάρχει μεγαλύτερη ακρίβεια, καθώς 1 μήνας (30 ημέρες) αντιστοιχεί περίπου σε 4,286 εβδομάδες. Επιπλέον, την ίδια υπολογιστική διαδικασία ακολουθεί και ο Klevmarken (2005). Έτσι, δημιουργήθηκε η μεταβλητή 'μηνιαίος μισθός' και κατ' επέκταση και η μεταβλητή 'μηνιαίες ώρες εργασίας', αφού προκειμένου να εξαχθεί η ελαστικότητα προσφοράς εργασίας ως προς το μισθό (δηλ. η ελαστικότητα μισθού) θα πρέπει οι δύο μεταβλητές να έχουν την ίδια μονάδα μέτρησης. Οι μηνιαίες ώρες εργασίας υπολογίστηκαν με τον ίδιο ακριβώς τρόπο που υπολογίστηκε και ο μηνιαίος μισθός, πολλαπλασιάζοντας δηλαδή τις δηλωμένες εβδομαδιαίες ώρες εργασίας με 4,3.

Από την άλλη μεριά, οι ερωτώμενοι που δήλωσαν αυτοαπασχολούμενοι, απαντούσαν στις εξής ερωτήσεις αναφορικά με τις ώρες εργασίας και τον μισθό:

*“πόσες ώρες την εβδομάδα, κατά μέσο όρο, εργαστήκατε τους τελευταίους 12 μήνες;” και “πόσα ήταν τα καθαρά προσωπικά σας έσοδα από την επιχείρησή σας για τους τελευταίους 12 μήνες; (δηλαδή μετά την καταβολή των φορολογικών και ασφαλιστικών εισφορών)”.*

Επομένως και οι αυτοαπασχολούμενοι (όπως και οι μισθωτοί) απαντούν σε εβδομαδιαία βάση για τις ώρες εργασίας. Όσον αφορά τον καθαρό μισθό όμως, απαντούν σε ετήσια βάση. Έτσι, για να μετατραπούν οι εβδομαδιαίες ώρες εργασίας σε μηνιαίες, ακολουθείται η ίδια διαδικασία που εφαρμόστηκε στους μισθωτούς, δηλαδή πολλαπλασιάζω με 4,3. Αντίθετα, για την μετατροπή των ετήσιων εσόδων σε μηνιαία, διαιρώ τα προσωπικά έσοδα από την επιχείρηση με το 12. Στη συνέχεια παίρνω αυτή τη μεταβλητή του 'μηνιαίου μισθού' των αυτοαπασχολούμενων και την συγχωνεύω με το 'μηνιαίο μισθό' των μισθωτών. Έτσι έχουμε τώρα τη μεταβλητή 'μηνιαίος μισθός' που αντικατοπτρίζει και τις δύο βασικές ομάδες του δείγματος (μισθωτούς και αυτοαπασχολούμενους). Το ίδιο ακριβώς πράγμα ισχύει και για τη μεταβλητή 'μηνιαίες ώρες εργασίας'. Από εδώ και στο εξής όποτε γίνεται αναφορά στο μισθό και τις ώρες εργασίας θα εννοείται μηνιαίος μισθός και μηνιαίες ώρες εργασίας.

Όσον αφορά το δείγμα μας τώρα, εκτός από τους άνεργους για τους οποίους φυσικά δεν υπάρχουν παρατηρήσεις σε αυτές τις δύο μεταβλητές, υπάρχουν και εργαζόμενοι που δεν έχουν δηλώσει είτε τις ώρες, είτε το μισθό, είτε και τα δύο. Πιο συγκεκριμένα από τους 418

εργαζόμενους (βλ. Πίνακα 3: μισθωτοί πλήρους και μερικής απασχόλησης + αυτοαπασχολούμενοι) δεν υπάρχει πληροφόρηση μόνο για 25 άτομα, όσον αφορά τις ώρες εργασίας, ενώ αναφορικά με το μισθό, οι αρνητικές τιμές φτάνουν τα 52 άτομα. Τα άτομα που δεν έχουν απαντήσει ούτε στο μισθό ούτε στις ώρες εργασίας είναι 5. Επίσης υπάρχουν δύο μισθωτοί, που ενώ έχουν απαντήσει και στο πόσες ώρες εργάζονται αλλά και στο μισθό που παίρνουν, εντούτοις δεν έχουν απαντήσει στην ερώτηση κάθε πότε πληρώνονται (εβδομαδιαία/δεκαπενθήμερη/μηνιαία βάση), με αποτέλεσμα να μην γνωρίζουμε σε τι βάση είναι ο δηλωμένος μισθός τους και ως επακόλουθο να μην είναι δυνατό να μετατραπεί στη μηνιαία βάση αναφοράς μας. Γι' αυτό το λόγο, μένουν εκτός δείγματος αυτές οι δύο παρατηρήσεις. Επίσης εκτός δείγματος μένουν και όσοι δεν απάντησαν είτε στις ώρες εργασίας είτε στο μισθό. Άρα το αρχικό δείγμα των 418 εργαζομένων γίνεται τώρα 344. Επομένως το τελικό μας δείγμα αποτελείται από 390 παρατηρήσεις (344 εργαζόμενοι + 46 άνεργοι).

### 5.2.2 Δευτερεύουσες μεταβλητές

Οι δευτερεύουσες μεταβλητές αφορούν κυρίως ατομικά χαρακτηριστικά και χαρακτηριστικά της εργασίας και είναι οι εξής: ηλικία, φύλο, εκπαίδευση, αριθμός ατόμων από τα οποία αποτελείται το νοικοκυριό, αριθμός ατόμων που συνεισφέρουν στο εισόδημα του νοικοκυριού, εργασιακή εμπειρία, απασχόληση στον ιδιωτικό ή στο δημόσιο τομέα, φύση της εργασίας, και τέλος μία ψευδομεταβλητή που δείχνει αν το άτομο θεωρεί τον εαυτό του ως άτομο καριέρας.

Η εκπαίδευση χρησιμοποιείται στη βιβλιογραφία με δύο τρόπους. Είτε ως συνεχής μεταβλητή που μετριέται σε χρόνια (έτη εκπαίδευσης), είτε ως κατηγορική μεταβλητή μετρημένη σε επίπεδα (επίπεδο εκπαίδευσης). Η συντριπτική πλειοψηφία πάντως, φαίνεται να έχει προτίμηση στα έτη εκπαίδευσης (βλέπε πχ, Daouli et al., 2004; Schaffner, 2002; Bishop et al., 2009; Hill, 1989; Leung and Yu, 2000; Parker et al., 2005; Co et al., 2005; Yamagata and Orme, 2005; Heckman and Macurdy, 1980; Cogan, 1981; Mroz, 1987, Takahashi et al., 2009). Ο λόγος που πιθανόν να το κάνουν αυτό, εκτός από το ότι υπάρχουν τα δεδομένα σε αυτή τη μορφή, είναι γιατί χρησιμοποιούν τη μεταβλητή εκπαίδευση καθώς και το τετράγωνο αυτής (βλ. Takahashi et al., 2009) και επιπλέον κάποιοι χρησιμοποιούν και όρους αλληλεπίδρασης (*interaction terms*) της εκπαίδευσης με άλλες μεταβλητές, όπως για παράδειγμα με την ηλικία (βλ. Yamagata and Orme, 2005; Mroz, 1987). Αντίθετα, λίγες είναι οι έρευνες που χρησιμοποιούν την εκπαίδευση σε επίπεδα, με τη βοήθεια ψευδομεταβλητών



(Nam, 1991; Kidd and Ferko, 2001; Dessing, 2002; Nawata and Ii, 2004; Wetzels, 2005; Donni and Moreau, 2007). Σε αυτήν τη περίπτωση, οι ψευδομεταβλητές είναι συνήθως τόσες όσες είναι και τα επίπεδα εκπαίδευσης, ή συμπυκνωμένες, περιλαμβάνοντας περισσότερες κατηγορίες και αντικατοπτρίζοντας χαμηλά, μεσαία και υψηλότερα επίπεδα εκπαίδευσης (Wetzels, 2005), ή πιο σπάνια, τα επίπεδα εκπαίδευσης εκφράζονται με μία μόνο ψευδομεταβλητή (Dessing, 2002). Στη παρούσα ανάλυση υπάρχουν 7 επίπεδα εκπαίδευσης, τα οποία συγκεντρώνονται σε τρεις κατηγορίες-ψευδομεταβλητές, βάσει της ISCED (International Standard Classification of Education, 1997). Οι τρεις αυτές ψευδομεταβλητές αφορούν το χαμηλό, μεσαίο και υψηλό επίπεδο εκπαίδευσης, όπως γίνεται και στην ανάλυση των del Campo et al. (2008). Στον Πίνακα 4, στο τέλος της ενότητας, αναλύεται ο διαχωρισμός των εκπαιδευτικών βαθμίδων στις τρεις κατηγορίες της “χαμηλής”, “μεσαίας” και “υψηλής” εκπαίδευσης, βάσει του ISCED 1997.

Η εργασιακή εμπειρία αφορά τα χρόνια που κάποιος είναι στην αγορά εργασίας και οι μέθοδοι μέτρησής της ποικίλουν επίσης. Συνήθως υπολογίζεται η δυνητική εμπειρία με βάση την ηλικία του ατόμου και τα έτη εκπαίδευσης. Για παράδειγμα, η Shaw (1992) υπολογίζει την εμπειρία αφαιρώντας από την ηλικία, τα έτη εκπαίδευσης και τον αριθμό 5 (ηλικία – έτη εκπαίδευσης – 5). Αντίστοιχα, οι Mizala et al. (1999) και οι Aaberge et al. (1999) την υπολογίζουν με τον ίδιο ακριβώς τρόπο, αφαιρώντας όμως τον αριθμό 6 αντί για 5, ενώ ο Maglad (1998) αφαιρεί τον αριθμό 7. Οι αριθμοί αυτοί που αφαιρούνται αντιστοιχούν στην ηλικία εισαγωγής στο σχολικό σύστημα. Οι μικρές διαφορές πιθανόν να οφείλονται στο αν λαμβάνεται υπόψη η προ-νηπιακή εκπαίδευση ή στη διαφορετική νομοθεσία που επικρατεί σε κάποια χώρα ως προς το σε ποια ηλικία υποχρεούται κάποιος να ξεκινήσει το σχολείο. Στη παρούσα μελέτη, για τον υπολογισμό της εμπειρίας ακολουθείται η μέθοδος που χρησιμοποιούν οι Co et al. (2005) και η El-Hamidi (2003). Δηλαδή η εμπειρία υπολογίζεται ως το χρονικό διάστημα από τότε που κάποιος μπήκε στην αγορά εργασίας για πρώτη φορά, μέχρι σήμερα (ως ‘σήμερα’ εννοείται η χρονολογία που διεξάγεται κάθε έρευνα ή γενικά, σε ποια χρονολογία αναφέρονται τα δεδομένα που χρησιμοποιεί ο ερευνητής στην ανάλυσή του). Ο λόγος που ακολουθείται αυτή η μέθοδος είναι γιατί υπάρχουν δεδομένα ως προς το ποια χρονιά ξεκίνησε να εργάζεται κάποιος για πρώτη φορά. Έτσι, δεδομένου ότι τα στοιχεία μας αφορούν το έτος 2004, η εμπειρία υπολογίζεται αφαιρώντας από το 2004, το έτος εισαγωγής στην αγορά εργασίας. Για παράδειγμα, για κάποιον που ξεκίνησε να εργάζεται το 1974, η εμπειρία του είναι 30 χρόνια.

Το είδος της εργασίας αφορά το αν η εργασία είναι χειρωνακτικής φύσεως ή όχι. Οι όροι που χρησιμοποιούνται στον κλάδο των οικονομικών της εργασίας γι’ αυτόν τον

διαχωρισμό είναι white-collar και blue-collar. White-collar είναι αυτός που δεν απασχολείται σε χειρωνακτικές θέσεις εργασίας, σε αντιδιαστολή με τον blue-collar που απασχολείται σε χειρωνακτικές θέσεις εργασίας<sup>15</sup>. Χαρακτηριστικό του white-collar είναι να εργάζεται σε ένα περιβάλλον με πολλά γραφεία και να έχει μπροστά του έναν υπολογιστή. Στη παρούσα μελέτη δημιουργήθηκε η ψευδομεταβλητή white collar με βάση μία ερώτηση στο ερωτηματολόγιο που αφορούσε τη κύρια απασχόληση. Ο ερωτώμενος καλούταν να απαντήσει ανάμεσα σε 9 κατηγορίες επαγγελμάτων για το ποιο είναι το κύριο επάγγελμά του. Η 5<sup>η</sup>, η 8<sup>η</sup> και η 9<sup>η</sup> κατηγορία θεωρήθηκαν ως επαγγέλματα blue-collar, ενώ οι υπόλοιπες ως white-collar. Στον Πίνακα 5 φαίνονται οι εννιά κατηγορίες επαγγελμάτων.

Συνοψίζοντας, η εκπαίδευση μαζί με την εργασιακή εμπειρία και τη φύση της εργασίας είναι μεταβλητές που κατασκευάστηκαν-δημιουργήθηκαν, ενώ οι υπόλοιπες μεταβλητές που χρησιμοποιούνται στη παρούσα μελέτη ως ανεξάρτητες μεταβλητές, είναι στην αρχική τους μορφή. Στον Πίνακα Α.1 του παραρτήματος Α, φαίνονται τα ονόματα που έχουν δοθεί στις μεταβλητές, οι ερμηνείες τους, καθώς και ποιες από αυτές είναι ψευδομεταβλητές.

**Πίνακας 4: Επίπεδα και κατηγορίες εκπαίδευσης**

Επίπεδο	Εκπαιδευτικός τίτλος	Κατηγορίες
1	Κανένας/Δεν έχω ολοκληρώσει κάποια εκπαίδευση	Χαμηλή εκπαίδευση
2	Δημοτικό	
3	Γυμνάσιο, Τριτάξιο	
4	Λύκειο (Γενικό, Ενιαίο, Τεχνικό), ΤΕΕ, Εξατάξιο Γυμνάσιο	Μεσαία εκπαίδευση
5	Τεχνικές Επαγγελματικές Σχολές (ΚΕΚ, ΙΕΚ)	
6	Τριτοβάθμια εκπαίδευση (ΤΕΙ/ΑΣΠΑΙΤΕ, ΑΕΙ, ΑΕΠ, Ανώτερες Σχολές)	Υψηλή εκπαίδευση
7	Μεταπτυχιακές σπουδές (Μάστερ, Διδακτορικό)	

<sup>15</sup> Οι θέσεις εργασίας που χαρακτηρίζονται ως μη χειρωνακτικές είναι πχ καθηγητές, δημόσιοι υπάλληλοι, διευθυντές, λογιστές, οικονομικοί και χρηματοοικονομικοί αναλυτές, προγραμματιστές, γιατροί. Αντίθετα, οι χειρωνακτικές θέσεις εργασίας περιλαμβάνουν κάθε είδους μηχανικούς και τεχνίτες, εργάτες που δουλεύουν σε εργοστάσια, οικοδομές, κατασκευές, εξορύξεις μεταλλευμάτων, καθώς και άτομα που εργάζονται στον πρωτογενή τομέα, όπως αγρότες, γεωργοί, κτηνοτρόφοι, ψαράδες, δασοκόμοι (βλ. LaRocco et al., 1980; Kauppinen-Toropainen et al., 1983; Groshen and Williams, 1992).

## Πίνακας 5: Κατηγορίες επαγγελμάτων

Κατηγορίες επαγγέλματος	Επάγγελμα
1	Διευθυντικά στελέχη-Μάνατζερς (π.χ. διευθυντικά και ανώτερα στελέχη σε ιδιωτικούς και δημόσιους οργανισμούς)
2	Άλλοι εξειδικευμένοι επαγγελματίες (π.χ. φυσικοί επιστήμονες, δάσκαλοι, επαγγελματίες στον τομέα της υγείας, δικηγόροι, κ.λπ.)
3	Υφιστάμενοι επαγγελματίες & εξειδικευμένο τεχνικό προσωπικό (π.χ. προγραμματιστές Η/Υ, επαγγελματίες στον τομέα της υγείας ή σε επιχειρήσεις σε υφιστάμενες θέσεις, κ.λπ.)
4	Γραφικά & γραμματειακά επαγγέλματα (π.χ. ταμίας, γραμματέας, δακτυλογράφος, τηλεφωνητής/τρια, αποθηκάριος, κ.λπ.)
5	Τεχνίτες & συναφή επαγγέλματα (π.χ. εργάτης/τρια σε τεχνική δουλειά όπως στις κατασκευές, στις ηλεκτροσυγκολήσεις, στην ξυλουργία, σε βιομηχανίες κατασκευής οχημάτων, κ.λπ.)
6	Παροχή προσωπικών υπηρεσιών & υπηρεσιών προστασίας (π.χ. εργαζόμενος/νη στις ένοπλες δυνάμεις, εταιρίες σεκιούριτι, ξεναγοί, κομμωτές/τριες, αισθητικοί, οικιακές δουλειές, κ.λπ.)
7	Πωλήσεις (π.χ. χρηματιστές, αντιπρόσωποι πωλήσεων & πωλητές/τριες, κ.λπ.)
8	Εργάτες σε βιομηχανίες & χειριστές μηχανημάτων (χειριστές μηχανημάτων π.χ. σε βιομηχανίες τροφίμων και καπνού, χημικών, πλαστικών, μετάλλου, σε οδικές μεταφορές, κ.λπ.)
9	Άλλα επαγγέλματα (π.χ. γεωργία, αλιεία, θήρα, κατασκευές, μεταφορές, επικοινωνίες, κ.λπ.)

## 5.3 Εμπειρική ανάλυση

### 5.3.1 Μοντέλο Heckman

Η διαδικασία που ακολουθείται για την εκτίμηση της συνάρτησης προσφοράς εργασίας υιοθετεί την προσέγγιση των Bishop et al. (2009) και Daouli et al. (2004). Αρχικός στόχος είναι να ελεγχθεί αν υπάρχει πρόβλημα μεροληψίας από την επιλογή του δείγματος (*sample selection bias*). Για να γίνει αυτό εφαρμόζεται το μοντέλο Heckman. Όπως είναι γνωστό (βλ. υποκεφάλαιο 4.2), το μοντέλο Heckman περιλαμβάνει δύο παλινδρομήσεις. Στο πρώτο στάδιο έχουμε την εξίσωση επιλογής (*selection equation*) και στο δεύτερο την εξίσωση αποτελέσματος (*outcome equation*). Η εξίσωση επιλογής εκτιμάται με μία probit παλινδρόμηση, ενώ η εξίσωση αποτελέσματος με μία ols παλινδρόμηση. Για να γίνει ταυτοποίηση του μοντέλου (βλ. υποκεφάλαιο 4.3), η εξίσωση επιλογής πρέπει να περιλαμβάνει όλες τις ανεξάρτητες μεταβλητές που χρησιμοποιούνται στην εξίσωση αποτελέσματος, και τουλάχιστον μία ακόμη που δεν επηρεάζει την εξίσωση αποτελέσματος

αλλά επηρεάζει την εξίσωση επιλογής (βλ. Sartori, 2003; Jones, 2007: chapter 7). Αυτή η μεταβλητή είναι γνωστή ως τεχνητή μεταβλητή (*instrument*).

Στη συγκεκριμένη μελέτη, η εξίσωση επιλογής είναι μία εξίσωση που προβλέπει τη πιθανότητα κάποιος να εργάζεται (δηλαδή τη πιθανότητα κάποιος να εισέλθει στην αγορά εργασίας). Εξαρτημένη μεταβλητή είναι μία ψευδομεταβλητή που παίρνει τιμή 1 αν κάποιος μέσα στο δείγμα εργάζεται και τιμή 0 αν είναι άνεργος. Από την άλλη μεριά, η εξίσωση αποτελέσματος είναι μία εξίσωση μισθού (*wage equation*). Δηλαδή εξαρτημένη μεταβλητή είναι ο μισθός (για την ακρίβεια ο φυσικός λογάριθμος του μισθού). Η μεταβλητή που χρησιμοποιείται ως *instrument* στην εξίσωση επιλογής είναι η ψευδομεταβλητή που δείχνει αν κάποιος θεωρεί τον εαυτό του ως άτομο καριέρας ή όχι. Ο λόγος που επιλέχθηκε αυτή η μεταβλητή ως *instrument* είναι γιατί θεωρητικά αν κάποιος θέλει και κυνηγάει να κάνει καριέρα σε ένα επάγγελμα, έχει πολύ μεγαλύτερες πιθανότητες να εισέλθει στην αγορά εργασίας, ενώ συγχρόνως αυτό δεν επηρεάζει αναγκαστικά και το μισθό που θα του προσφέρουν οι εργοδότες. Παράλληλα η μεταβλητή αυτή, όταν εισάγεται ως ανεξάρτητη στην εξίσωση μισθού δεν είναι στατιστικά σημαντική.

Στον Πίνακα 6 φαίνονται τα περιγραφικά χαρακτηριστικά όλων των μεταβλητών (μέση τιμή και τυπική απόκλιση) για όλο το δείγμα, αλλά και ξεχωριστά για τους εργαζόμενους και τους άνεργους. Το μέγεθος του δείγματος βάσει όλων των ανεξάρτητων μεταβλητών φτάνει τις 358 παρατηρήσεις. Η μέση ηλικία του δείγματος είναι τα 52 χρόνια, ενώ η εργασιακή εμπειρία κατά μέσο όρο είναι γύρω στα 31 χρόνια. Το 65% του δείγματος αποτελείται από άντρες και το υπόλοιπο 35% από γυναίκες. Όσον αφορά την εκπαίδευση, φαίνεται ότι τα άτομα του δείγματος είναι ομοιόμορφα κατανομημένοι στις τρεις κατηγορίες, με το 30% να ανήκουν στη μεσαία εκπαίδευση. Επίσης, κάθε νοικοκυριό αποτελείται κατά μέσο όρο από 3 άτομα, ενώ τα άτομα που συνεισφέρουν στο οικογενειακό εισόδημα είναι περίπου 2. Τέλος, βλέπουμε ότι το 72% εργάζονται στον ιδιωτικό τομέα, το 62% απασχολούνται σε μη χειρωνακτικές θέσεις εργασίας και το 28% δηλώνει ότι είναι άτομο καριέρας. Αναφορικά με τους άνεργους, βλέπουμε ότι τα ποσοστά είναι αναμενόμενα, με τους περισσότερους (63%) να έχουν χαμηλή εκπαίδευση, μόλις το 5% να είναι άτομα καριέρας, και το 95% να εργάζονταν στον ιδιωτικό τομέα προτού μείνουν άνεργοι.

Πίνακας 6: Περιγραφικά χαρακτηριστικά (μέση τιμή και τυπική απόκλιση)

	Συνολικό δείγμα	Εργαζόμενοι	Άνεργοι
Αριθμός παρατηρήσεων	358	318	40
hours		181,19 (56)	
wage		1425,75 (990,86)	
age	51,78 (5,38)	51,74 (5,40)	52,05 (5,25)
gender	0,65 (0,48)	0,68 (0,47)	0,43 (0,50)
laborexpr	31,29 (7,50)	31,10 (7,44)	32,85 (7,96)
loweduc	0,39 (0,49)	0,36 (0,48)	0,63 (0,49)
mideduc	0,30 (0,46)	0,30 (0,46)	0,27 (0,45)
higheduc	0,31 (0,46)	0,34 (0,47)	0,10 (0,30)
privsector	0,72 (0,45)	0,69 (0,47)	0,95 (0,22)
housemembers	3,38 (1,30)	3,34 (1,30)	3,75 (1,26)
whitecollar	0,62 (0,48)	0,64 (0,48)	0,43 (0,50)
personscontrib	1,68 (0,71)	1,71 (0,72)	1,43 (0,55)
careerperson	0,28 (0,45)	0,31 (0,47)	0,05 (0,22)

Notes: σε παρένθεση η τυπική απόκλιση

Όπως διαπιστώθηκε στα υποκεφάλαια 4.3 και 4.4, το μοντέλο του Heckman μπορεί να υπολογιστεί με δύο τρόπους. Ο ένας είναι ο τυπικός τρόπος των δύο βημάτων (*two-step procedure*) ο οποίος συχνά αναφέρεται ως Heckit και ο άλλος είναι με τη μέθοδο μεγίστης πιθανοφάνειας. Τα υπέρ και τα κατά των δύο τρόπων παρουσιάστηκαν στο κεφάλαιο 4. Στη συγκεκριμένη μελέτη εφαρμόζονται και οι δύο τρόποι προκειμένου να φανούν τυχόν διαφορές στα αποτελέσματα. Οι δύο εξισώσεις παλινδρόμησης είναι οι εξής:

$$\text{Selection equation: Participation} = \alpha_0 + \alpha_1 Z_i + u_i \quad (5.3.1)$$

$$\text{Wage equation: } \ln \text{wage} = \beta_0 + \beta_1 X_i + \beta_2 \lambda(\alpha_1 Z_i) + \varepsilon_i \quad (5.3.2)$$

όπου  $X_i$  είναι το διάνυσμα των ανεξάρτητων μεταβλητών,  $Z_i$  το διάνυσμα που περιλαμβάνει ό,τι περιλαμβάνει το  $X_i$  και επιπλέον τη μεταβλητή instrument,  $\lambda(\alpha_1 Z_i) = \frac{\varphi(\alpha_1 Z_i)}{\Phi(\alpha_1 Z_i)}$  είναι το inverse mill's ratio,  $\beta_2 = \rho\sigma_\varepsilon$ , και  $U_i, \varepsilon_i$  είναι οι διαταρακτικοί όροι. Οι υποθέσεις που ισχύουν για το μοντέλο του Heckman με τη μέθοδο των δύο βημάτων είναι ότι τα σφάλματα  $u_i$  ακολουθούν την τυπική κανονική κατανομή και  $u_i, \varepsilon_i$  ανεξάρτητα από τις ερμηνευτικές μεταβλητές. Αντίθετα η μέθοδος μεγίστης πιθανοφάνειας στηρίζεται σε πιο αυστηρές υποθέσεις, ότι δηλαδή τα δύο σφάλματα ακολουθούν μία κοινή κανονική κατανομή  $(u_i, \varepsilon_i) \sim \left(0, \begin{bmatrix} 1 & \rho\sigma_\varepsilon \\ \rho\sigma_\varepsilon & \sigma_\varepsilon^2 \end{bmatrix}\right)$  (bivariate normal distribution) και  $u_i \sim N(0,1), \varepsilon_i \sim N(0, \sigma_\varepsilon^2)$ . Στον Πίνακα 7 παρουσιάζονται τα αποτελέσματα του μοντέλου Heckman με τη διαδικασία των δύο βημάτων και στον Πίνακα 8 τα αντίστοιχα αποτελέσματα με τη μέθοδο μεγίστης πιθανοφάνειας.

Συγκρίνοντας τους δύο πίνακες βλέπουμε ότι δεν υπάρχουν μεγάλες διαφορές στις εκτιμήσεις. Ειδικά οι εκτιμήσεις της εξίσωσης του μισθού είναι σχεδόν ίδιες. Αυτό που παρατηρεί κανείς είναι η αλλαγή των προσήμων σε κάποιες μεταβλητές. Για παράδειγμα, αναφορικά με την εξίσωση μισθού, βλέπουμε ότι ενώ με τη διαδικασία δύο βημάτων προκύπτει αρνητικό πρόσημο στη μεταβλητή φύλο, με τη μέθοδο μεγίστης πιθανοφάνειας προκύπτει θετικό. Αντίστοιχα, αναφορικά με την εξίσωση επιλογής, οι μεταβλητές με αντίθετα πρόσημα είναι η ηλικία και η μεσαία εκπαίδευση. Από τις υπόλοιπες μεταβλητές που δεν υπάρχει πρόβλημα με αντίθετα πρόσημα ανάμεσα στις δύο μεθόδους εκτίμησης, τα πρόσημα που αφορούν τον αριθμό των ατόμων στο νοικοκυριό καθώς και τα άτομα που κυνηγούν τη καριέρα είναι τα αναμενόμενα. Ειδικά στη ψευδομεταβλητή “careerperson” πρέπει να δοθεί ιδιαίτερη βάση καθώς είναι το instrument και τυχόν αρνητικό πρόσημο θα δημιουργούσε τεράστιο πρόβλημα. Εκτός αυτού είναι αυτονόητο ότι κάποιος που θεωρεί τον εαυτό του ως άτομο καριέρας έχει μεγαλύτερες πιθανότητες να συμμετέχει στην αγορά εργασίας, αφού εκ φύσεως είναι πρακτικά αδύνατο να κάνει κάποιος καριέρα μένοντας εκτός εργασίας. Παρατηρούμε ότι η ψευδομεταβλητή αυτή είναι στατιστικά σημαντική σε επίπεδο στατιστικής σημαντικότητας 0,05 με τη μέθοδο εκτίμησης δύο βημάτων, ενώ με τη μέθοδο μεγίστης πιθανοφάνειας είναι στατιστικά σημαντική σε επίπεδο 0,01. Αντίθετα, για τη μεταβλητή personscontrib ίσως θα περιμέναμε εκ των προτέρων να έχει αρνητικό πρόσημο, αφού λογικό είναι να σκεφτεί κανείς ότι όσα περισσότερα άτομα συνεισφέρουν στο οικογενειακό εισόδημα, τόσο θα ελαττώνεται η πιθανότητα να εισέλθει στην αγορά εργασίας, με την έννοια δηλαδή ότι μπορεί να στηρίζεται στους άλλους. Στη συγκεκριμένη περίπτωση το θετικό πρόσημο πιθανόν να βασίζεται σε κάποιους ‘κρυφούς’, μη μετρήσιμους και μη

παρατηρήσιμους προσωπικούς παράγοντες, όπως για παράδειγμα ότι ψυχολογικά δεν νιώθει καλά ένα άτομο όταν βλέπει να συνεισφέρουν όλοι οι άλλοι στο εισόδημα και αυτός να κάθεται να το απολαμβάνει χωρίς να συνεισφέρει καθόλου. Άρα υπό αυτό το πρίσμα, όσα περισσότερα άτομα συνεισφέρουν στο εισόδημα, τόσο περισσότερο θα αυξάνεται η πιθανότητα για κάποιον να εργαστεί, προκειμένου να συνεισφέρει και αυτός ένα μέρος του εισοδήματος.

**Πίνακας 7: Heckman two-step procedure**

Indep. variables	Heckman two-step model	
	Selection equation (Probit)	Wage equation (ols)
age	-0,006 (0,028)	-0,011 (0,010)
gender	0,874*** (0,250)	-0,008 (0,119)
laborexp	-0,023 (0,021)	0,011 (0,008)
loweduc	-1,615 (0,329)	-0,444*** (0,116)
mideduc	-0,054 (0,338)	-0,223** (0,107)
housemembers	-0,214** (0,084)	0,064* (0,036)
privsector	-1,102*** (0,383)	-0,217** (0,108)
whitecollar	0,379 (0,246)	-0,174* (0,103)
personscontrib	0,540*** (0,180)	-0,259*** (0,072)
careerperson	0,800** (0,343)	- -
cons.	2,37* (1,304)	8,093*** (0,476)
mills ratio (lambda)		-0,704* (0,365)
rho		-0,98
Pseudo R-squared	0,244	-
LR Chi-sq (10)	61,11	-
Wald Chi-sq (18)	-	74,49
P-value	0,000	0,000
N	358	358

Notes: \*\*\*significant at 1%, \*\*at 5%, \*at 10% (σε παρένθεση τα τυπικά σφάλματα)

Πίνακας 8: Heckman ML estimation

Indep. variables	Heckman MLE	
	Selection equation	Wage equation
age	0,002 (0,030)	-0,011 (0,011)
gender	0,837*** (0,249)	0,009 (0,091)
laborexp	-0,032 (0,024)	0,010 (0,008)
loweduc	-0,031 (0,305)	-0,451*** (0,113)
mideduc	0,024 (0,304)	-0,225** (0,102)
housemembers	-0,237*** (0,073)	0,058 (0,035)
privsector	-0,918*** (0,323)	-0,234*** (0,083)
whitecollar	0,204 (0,235)	-0,176* (0,095)
personscontrib	0,458*** (0,161)	-0,251*** (0,059)
careerperson	1,074*** (0,332)	- -
cons.	2,246 (1,368)	8,106*** (0,501)
mills ratio (lambda)		-0,601*** (0,072)
rho		-0,86
Log-likelihood		-400,28
Wald Chi-sq (9)		54,05
P-value		0,000
N		358

Notes: \*\*\*significant at 1%, \*\*at 5%, \*at 10% (σε παρένθεση τα robust τυπικά σφάλματα)

Το πιο σημαντικό στοιχείο των δύο πινάκων, είναι το inverse mill's ratio. Αυτός ο 'λόγος' είναι ουσιαστικά που θα μας δείξει αν υπάρχει selection bias. Μία στατιστικά σημαντική εκτίμηση σημαίνει ότι υπάρχει πρόβλημα (βλ. υποκεφάλαιο 4.4). Κοιτώντας τον Πίνακα 7, βλέπουμε καταρχήν ότι ο συντελεστής συσχέτισης των καταλοίπων των δύο παλινδρομήσεων-εξισώσεων είναι πολύ υψηλός ( $\rho = -0,98$ ), παρόλο που ο λόγος inverse mill's είναι στατιστικά σημαντικός μόνο σε επίπεδο στατιστικής σημαντικότητας 10%. Αυτός είναι ένας παραπάνω λόγος που εκτιμήθηκε το μοντέλο του Heckman και με τη μέθοδο ML, καθώς με τη μέθοδο δύο βημάτων δεν ήταν εντελώς ξεκάθαρο για το αν υπάρχει πρόβλημα. Πηγαίνοντας λοιπόν στον Πίνακα 8, βλέπουμε ότι τώρα ο λόγος είναι στατιστικά σημαντικός



σε επίπεδο 1% και ο συντελεστής συσχέτισης  $\rho$  παραμένει υψηλός. Οπότε, φαίνεται να υπάρχει πρόβλημα. Όπως είδαμε όμως στο υποκεφάλαιο 4.4, ο έλεγχος στατιστικής σημαντικότητας με το t-test δεν είναι αξιόπιστος αν υπάρχει πολυσυγγραμμικότητα (Yamagata and Orme, 2005). Γι' αυτό το λόγο πραγματοποιήθηκαν και δύο επιπλέον έλεγχοι. Οι έλεγχοι αυτοί είναι τα τεστ LR και Wald, όπως προκύπτουν από την εκτίμηση μεγίστης πιθανοφάνειας. Σύμφωνα με τους Nawata & McAleer (2001) το τεστ LR είναι εύρωστο ακόμη και όταν υπάρχει πολυσυγγραμμικότητα, σε αντίθεση με το Wald τεστ. Συνολικά οι τρεις έλεγχοι φαίνονται στον Πίνακα 9. Τα τεστ Wald και LR ελέγχουν τη μηδενική υπόθεση περί ανεξαρτησίας μεταξύ των δύο εξισώσεων (εξίσωση επιλογής και εξίσωση μισθού). Η μηδενική υπόθεση δηλαδή είναι  $\rho=0$  (όπως και στο t-test), όπου  $\rho$  είναι ο συντελεστής συσχέτισης των καταλοίπων  $u_i$  και  $\varepsilon_i$  των δύο εξισώσεων (βλ. εξίσωση 5.3.1 και 5.3.2). Όπως είδαμε προηγουμένως, με το t-test απορρίπτουμε τη μηδενική υπόθεση σε επίπεδο 10%, ενώ βλέπουμε ότι με τα τεστ Wald και LR την απορρίπτουμε σε όλα τα επίπεδα στατιστικής σημαντικότητας.

**Πίνακας 9: Έλεγχοι υποθέσεων**

	P-value
t (two-step)	0,054
Wald	0,0000
LR	0,0006

Το επόμενο βήμα είναι ο έλεγχος για πολυσυγγραμμικότητα, προκειμένου να φανεί πιο από τα τρία παραπάνω τεστ ελέγχου είναι πιο αξιόπιστο, και ως επακόλουθο και ποια μέθοδος εκτίμησης του μοντέλου Heckman είναι περισσότερο αξιόπιστη. Σύμφωνα με τους Nawata & Nagase (1996), η εκτίμηση μεγίστης πιθανοφάνειας αποδίδει καλύτερα από την εκτίμηση των δύο βημάτων όταν υπάρχει πολυσυγγραμμικότητα. Για να ελεγχθεί η πολυσυγγραμμικότητα, ακολουθείται η μέθοδος της βοηθητικής *ols* παλινδρόμησης (βλ. υποκεφάλαιο 4.4). Αυτή η παλινδρόμηση έχει ως εξαρτημένη μεταβλητή το *inverse mill's ratio*, έτσι όπως εκτιμήθηκε στο πρώτο βήμα, και ως ανεξάρτητες το διάνυσμα  $X_i$  (Nelson, 1984; Leung and Yu, 1996; Yamagata and Orme, 2005). Ο (προσαρμοσμένος) συντελεστής προσδιορισμού  $R^2$  είναι αυτός που θα δείξει αν υπάρχει πολυσυγγραμμικότητα. Όσο μεγαλύτερος ο συντελεστής, τόσο μεγαλύτερο και το πρόβλημα της πολυσυγγραμμικότητας. Στον Πίνακα A.2 του παραρτήματος είναι τα αποτελέσματα της βοηθητικής παλινδρόμησης. Στη μία στήλη εξαρτημένη είναι το *inverse mill's ratio* όπως εκτιμήθηκε με το μοντέλο δύο βημάτων, και στην άλλη στήλη το αντίστοιχο, όπως εκτιμήθηκε με τη μέθοδο μεγίστης πιθανοφάνειας. Ο

προσαρμοσμένος συντελεστής προσδιορισμού είναι περίπου 72% στη πρώτη περίπτωση και 68% στη δεύτερη. Δηλαδή δεν διαφέρουν και πολύ μεταξύ τους. Επομένως ίσως να υπάρχει πολυσυγγραμμικότητα, αλλά σε πολύ μικρό βαθμό, αφού ο συντελεστής προσδιορισμού δεν είναι ιδιαίτερα υψηλός.

Συνοψίζοντας, είδαμε ότι το μοντέλο του Heckman εκτιμήθηκε και με τις δύο μεθόδους εκτίμησης, και έδειξαν και οι δύο ότι υπάρχει selection bias στο υπόδειγμα. Στη συνέχεια έγιναν επιπλέον έλεγχοι για επιβεβαίωση. Τέλος έγινε έλεγχος και για πολυσυγγραμμικότητα, προκειμένου να διαπιστωθεί ποια από τις δύο μεθόδους εκτίμησης είναι πιο αξιόπιστη. Οι συντελεστές προσδιορισμού της βοηθητικής παλινδρόμησης είναι παρόμοιοι για τις δύο μεθόδους, επομένως φαίνεται ότι και οι δύο μέθοδοι στη συγκεκριμένη περίπτωση είναι ισοδύναμοι. Με λίγα λόγια, δεν έχει μεγάλη σημασία ποια από τις δύο μεθόδους θα υιοθετηθεί στο επόμενο κομμάτι της ανάλυσης, το οποίο είναι η εκτίμηση της συνάρτησης προσφοράς εργασίας.

### 5.3.2 Ελαστικότητα προσφοράς εργασίας

Η ελαστικότητα προσφοράς εργασίας προκύπτει από την εκτίμηση της συνάρτησης προσφοράς εργασίας. Η συνάρτηση προσφοράς εργασίας έχει ως εξαρτημένη μεταβλητή τις ώρες εργασίας και ως βασική ερμηνευτική μεταβλητή το μισθό. Η εξίσωση προσφοράς εργασίας θα εκτιμηθεί με το μοντέλο tobit, καθώς δεν υπάρχουν παρατηρήσεις των ωρών εργασίας για τους άνεργους. Αντίθετα, υπάρχουν πλέον οι παρατηρήσεις για το μισθό, έτσι όπως προκύπτουν από την εξίσωση μισθού του μοντέλου Heckman. Δηλαδή, για την εκτίμηση της εξίσωσης προσφοράς εργασίας δεν θα έχουμε τις πραγματικές τιμές του μισθού ως ανεξάρτητη, αλλά τις εκτιμημένες-προβλεπόμενες τιμές του μισθού. Με λίγα λόγια, με τη βοήθεια του μοντέλου Heckman εκτιμήθηκε ο μισθός που κάποιος θα έπαιρνε αν εργαζόταν, και ως επακόλουθο έχουμε πλέον θετικές τιμές και για τους άνεργους. Οι υπόλοιπες ανεξάρτητες μεταβλητές είναι κάποιες από αυτές που χρησιμοποιήθηκαν στο μοντέλο του Heckman. Ο λόγος που δεν χρησιμοποιούνται όλες οι ανεξάρτητες μεταβλητές, είναι γιατί δεν βγαίνουν στατιστικά σημαντικές (δηλαδή δεν επηρεάζουν τις ώρες εργασίας) και επιπλέον θα υπήρχε πρόβλημα τέλειας πολυσυγγραμμικότητας με τη μεταβλητή του εκτιμημένου μισθού, αφού ο εκτιμημένος μισθός εκτιμήθηκε με βάση όλες αυτές τις ανεξάρτητες μεταβλητές (El-Hamidi, 2003). Επομένως η εξίσωση προσφοράς εργασίας έχει την εξής μορφή:

$$\text{hours} = \gamma_0 + \gamma_1 \ln \text{wagehat} + \gamma_2 M_i + V_i \quad (5.3.3)$$

όπου  $\ln wage_{it}$  ο εκτιμημένος λογάριθμος του μισθού ( $\widehat{\ln wage}$ ),  $M_i$  το διάνυσμα των ανεξάρτητων μεταβλητών που περιλαμβάνει το instrument και κάποιες από τις μεταβλητές του διανύσματος  $X_i$  (βλ. εξισώσεις 5.3.1 και 5.3.2), και  $V_i$  ο διαταρακτικός όρος [ $V_i$  i.i.d και  $V_i \sim N(0, \sigma^2)$ ]. Οι ώρες εργασίας παίρνουν μηδενικές τιμές για τους άνεργους, αφού δεν μπορούν να είναι ποτέ μικρότερες από μηδέν (Birch, 2005). Για λόγους σύγκρισης και αντιστοιχίας με την προηγούμενη ενότητα, το tobit εκτιμήθηκε δύο φορές. Στη 1<sup>η</sup> περίπτωση ως ανεξάρτητη μεταβλητή είναι ο εκτιμημένος μισθός όπως εκτιμήθηκε με το μοντέλο Heckman δύο βημάτων, και στη δεύτερη περίπτωση αυτός που εκτιμήθηκε με το μοντέλο Heckman μεγίστης πιθανοφάνειας. Στον Πίνακα 10 φαίνονται τα αποτελέσματα του μοντέλου tobit και για τις δύο περιπτώσεις. Επίσης, στον Πίνακα Α.3 του παραρτήματος Α υπάρχει ο πίνακας με τους συντελεστές συσχέτισης όλων των ερμηνευτικών μεταβλητών που χρησιμοποιούνται στην εξίσωση προσφοράς εργασίας, προκειμένου να φανεί αν υπάρχει κάποια μεταβλητή που να δημιουργεί πρόβλημα πολυσυγγραμμικότητας στη εκτίμηση του μοντέλου tobit.

Όπως θα περίμενε κανείς κοιτώντας τον Πίνακα 10, τα αποτελέσματα δεν διαφέρουν πολύ, αφού όπως φάνηκε στη προηγούμενη ενότητα και οι δύο τρόποι εκτίμησης του μοντέλου Heckman είναι αξιόπιστοι. Επίσης, δεν υπάρχει πρόβλημα ούτε με τα πρόσημα, καθώς και οι δύο τρόποι δίνουν τα ίδια. Επιπλέον, οι συντελεστές συσχέτισης μεταξύ των ανεξάρτητων μεταβλητών είναι πολύ χαμηλοί (Πίνακας Α.3). Ο υψηλότερος συντελεστής συσχέτισης είναι αυτός μεταξύ της ηλικίας και της εργασιακής εμπειρίας ( $\rho=0,69$ ). Δεν είναι όμως τόσο υψηλός ώστε να δημιουργεί πρόβλημα πολυσυγγραμμικότητας η μία εκ των δύο μεταβλητών. Επειδή το μοντέλο προσφοράς εργασίας είναι γραμμικό-λογαριθμικό (δηλαδή ο μισθός είναι σε λογάριθμο, ενώ οι ώρες εργασίας δεν είναι) η ελαστικότητα προσφοράς εργασίας θα υπολογιστεί με βάση τη μέση τιμή των ωρών εργασίας (βλ. Χάλκος, 2011: σελ. 191). Δηλαδή η ελαστικότητα θα είναι  $\hat{\gamma}_1 / \overline{hours}$ . Ο συντελεστής  $\hat{\gamma}_1$  ισούται περίπου με -48,09 στη μία περίπτωση και -47,75 στη δεύτερη (βλ. Πίνακα 10). Η μέση τιμή των ωρών εργασίας με βάση τους εργαζόμενους είναι 181,19 ώρες/μήνα (βλ. Πίνακα 6). Άρα η ελαστικότητα είναι  $-48,09 / 181,19 = -0,265$  και  $-47,75 / 181,19 = -0,263$  αντίστοιχα για τις δύο περιπτώσεις του Heckman. Εν ολίγοις, η ελαστικότητα προσφοράς εργασίας υπολογισμένη στη μέση τιμή της εξαρτημένης μεταβλητής είναι -0,26. Το αρνητικό πρόσημο υποδηλώνει ότι επικρατεί το αποτέλεσμα εισοδήματος σε σχέση με το αποτέλεσμα υποκατάστασης. Δηλαδή μία αύξηση του λογαρίθμου του μισθού κατά 1 μονάδα έχει ως αποτέλεσμα τη μείωση των ωρών εργασίας κατά 0,26 ώρες. Επίσης κοιτώντας τις άλλες μεταβλητές, φαίνεται πως οι άνδρες εργάζονται περίπου 46 ώρες το μήνα περισσότερο από

τις γυναίκες (*ceteris paribus*), ενώ κάποιος που θεωρεί τον εαυτό του ως άτομο καριέρας εργάζεται περίπου 35 ώρες το μήνα περισσότερο (*ceteris paribus*) συγκριτικά με κάποιον που δεν τον ενδιαφέρει η επαγγελματική καριέρα.

Στον Πίνακα 10, στις δύο τελευταίες στήλες, είναι η συνάρτηση προσφοράς εργασίας εκτιμημένη με τη μέθοδο ελαχίστων τετραγώνων και για τις δύο μεθόδους εκτίμησης του μοντέλου Heckman, σε αντιστοιχία με το tobit. Εξαρτημένη μεταβλητή είναι οι ώρες εργασίας βάζοντας τη τιμή 0 σε όσους είναι άνεργοι. Οι δύο εκτιμήσεις ols δεν έχουν διαφορές μεταξύ τους, αλλά διαφέρουν αρκετά συγκρίνοντας τις με τις αντίστοιχες εκτιμήσεις του tobit. Οι μεγαλύτερες διαφορές παρατηρούνται στην εκπαίδευση, το φύλο και φυσικά στο μισθό. Ενώ με το tobit ο εκτιμημένος συντελεστής του μισθού είναι περίπου -48, με την ols είναι περίπου -41. Άρα η ελαστικότητα προσφοράς εργασίας με βάση το μοντέλο ols είναι  $-41 / 181,19 = -0,226 \approx -0,23$ . Σε σύγκριση με την ελαστικότητα του tobit η διαφορά δεν είναι και πολύ μεγάλη (-0,23 έναντι -0,26), αλλά μεταξύ των μοντέλων tobit και ols, προτιμότερο είναι το tobit, αφού λαμβάνει υπόψη ότι το δείγμα είναι λογοκριμένο. Αντίθετα, η μέθοδος ols θεωρεί ότι το δείγμα είναι ολόκληρο, με τιμές 0 ως παρατηρημένες τιμές των ωρών εργασίας για τους άνεργους (εφόσον ένας άνεργος δεν δουλεύει, έχει μηδενικές ώρες εργασίας).

Κατά κοινή ομολογία οι ελαστικότητες προσφοράς εργασίας των γυναικών έχουν θετικό πρόσημο, ενώ οι ελαστικότητες προσφοράς εργασίας των αντρών ποικίλουν. Μπορεί να είναι μεγάλες και θετικές, μεγάλες και αρνητικές ή και μηδενικές (Botjas, 2003: σελ. 76). Στο Παράρτημα Β υπάρχουν δύο πίνακες (Πίνακας Β.1 και Β.2) με συγκεντρωτικές ελαστικότητες από διάφορες έρευνες, έτσι όπως τις συγκεντρώσαν οι Bargain et al. (2011). Στους δύο αυτούς πίνακες υπάρχει το εύρος τιμών μέσα στο οποίο κυμαίνονται οι ελαστικότητες, και σε κάποιες περιπτώσεις υπάρχει η ακριβής εκτίμηση των ελαστικοτήτων. Όπως φαίνεται και από τους πίνακες, όλες οι ελαστικότητες των γυναικών έχουν θετικό πρόσημο, ενώ αντίθετα κάποιες ελαστικότητες των αντρών έχουν αρνητικό πρόσημο, με τις περισσότερες να κυμαίνονται πολύ κοντά στο μηδέν. Εκτός αυτού, όλες οι ελαστικότητες των αντρών είναι ανελαστικές, ενώ των γυναικών είναι περισσότερο ελαστικές. Αυτό που παρατηρούμε επίσης είναι ότι στην Ευρώπη οι ελαστικότητες των αντρών και γυναικών είναι πολύ κοντά μεταξύ τους (βλ. Πίνακα Β.1), δηλαδή δεν υπάρχουν μεγάλες διαφορές μεταξύ αντρών και γυναικών. Αντίθετα στις ΗΠΑ παρατηρούνται διαφορές στη προσφορά εργασίας μεταξύ αντρών και γυναικών (βλ. Πίνακα Β.2).

Σύμφωνα με αυτά λοιπόν, η εκτιμημένη ελαστικότητα της παρούσας μελέτης (-0,26) φαίνεται να συμφωνεί με τη βιβλιογραφία, αφού το μεγαλύτερο ποσοστό του δείγματος

(65%) αποτελείται από άντρες, οπότε το αρνητικό πρόσημο κατά πάσα πιθανότητα να οφείλεται σε αυτούς. Επιπλέον η ελαστικότητα είναι αρκετά ανελαστική, όπως προβλέπει η θεωρία για τους άντρες, αλλά και σε κάποιες περιπτώσεις και για τις γυναίκες (βλ. Πίνακα B.1).

**Πίνακας 10: Εκτίμηση συνάρτησης προσφοράς εργασίας**

Indep. variables	Tobit model		OLS model	
	Two-step Heckman prediction	MLE Heckman prediction	Two-step Heckman prediction	MLE Heckman prediction
lnwagehat	-48,095** (20,734)	-47,754** (20,947)	-40,89** (18)	-41,01** (18,02)
age	-1,416 (1,183)	-1,385 (1,181)	-1,38 (1,04)	-1,35 (1,04)
laborexp	0,457 (0,909)	0,399 (0,905)	0,48 (0,85)	0,43 (0,84)
gender	45,782*** (9,918)	46,538*** (10,036)	40,43*** (8,92)	41,14*** (9,03)
loweduc	-18,273 (13,554)	-18,746 (13,710)	-12,38 (12,25)	-12,91 (12,35)
mideduc	1,391 (12,595)	1,245 (12,648)	3,88 (11,31)	3,67 (11,34)
careerperson	34,581*** (10,372)	34,599*** (10,373)	31,85*** (9,55)	31,86*** (9,54)
cons.	527,183*** (162,043)	523,780*** (163,215)	477,97*** (134,94)	478,12*** (134,64)
Pseudo R-squared	0,011	0,011	-	-
R-squared			0,11	0,11
LR Chi-sq (7)	41,20	41,02	-	-
P-value	0,000	0,000	-	-
N	358	358	358	358

Notes: \*\*\*significant at 1%, \*\*at 5% (σε παρένθεση τα robust τυπικά σφάλματα)

Παρακάτω εκτιμάται η συνάρτηση προσφοράς εργασίας πάλι με τη μέθοδο ελαχίστων τετραγώνων. Μόνο που τώρα η ols παλινδρόμηση αφορά μόνο τα άτομα που εργάζονται και ως επιπρόσθετη ανεξάρτητη μεταβλητή χρησιμοποιείται το inverse mill's ratio. Τη μέθοδο αυτή ακολουθούν οι Bishop et al. (2009). Ακολουθώντας τη μέχρι τώρα διαδικασία, θα εκτιμηθούν δύο μοντέλα ols. Το πρώτο μοντέλο έχει ως ανεξάρτητες τις εκτιμημένες τιμές του λογαρίθμου του μισθού και το αντίστοιχο mill's ratio, όπως προέκυψαν από το heckman δύο βημάτων. Το δεύτερο μοντέλο έχει τις εκτιμημένες τιμές του λογαρίθμου του μισθού και το αντίστοιχο mill's ratio, έτσι όπως προέκυψαν από το heckman μεγίστης πιθανοφάνειας. Τα

αποτελέσματα φαίνονται στον Πίνακα 11. Παρατηρούμε πάλι ότι τα δύο μοντέλα μοιάζουν αρκετά μεταξύ τους, με μόνο την εκτίμηση του mill's ratio να διαφέρει αρκετά. Στη πρώτη περίπτωση μάλιστα είναι στατιστικά σημαντικός σε επίπεδο 10%, ενώ στη δεύτερη δεν είναι στατιστικά σημαντικός. Οι συντελεστές του λογαρίθμου του μισθού είναι επίσης στατιστικά μη σημαντικοί, ενώ ο συντελεστής προσδιορισμού είναι πολύ χαμηλός. Με λίγα λόγια η *ols* παλινδρόμηση της προσφοράς εργασίας μόνο για τους εργαζόμενους, με βάση τη διόρθωση για selection bias (συντελεστής mill's ratio), δεν δείχνει αξιόπιστη. Τι θα γίνει όμως αν δεν συμπεριληφθεί το inverse mill's ratio και αντί για τις εκτιμημένες τιμές του μισθού, έχουμε τις πραγματικές; Αυτό φαίνεται στον Πίνακα 12.

Στον Πίνακα 12, είναι η *ols* παλινδρόμηση μόνο για τους εργαζόμενους, χωρίς τη διόρθωση του selection bias. Δηλαδή ως ανεξάρτητη δεν περιλαμβάνει τις εκτιμημένες τιμές του μισθού, αλλά τις πραγματικές, και δεν υπάρχει πλέον ούτε το mill's ratio. Με λίγα λόγια, είναι η παλινδρόμηση που τρέχουμε στα αρχικά δεδομένα, αν αφήσουμε εκτός δείγματος τους άνεργους και δεν ελέγχουμε για selection bias. Τώρα ως ανεξάρτητη μεταβλητή δεν έχουμε τις εκτιμημένες τιμές του λογαρίθμου του μισθού ( $\ln wage_{hat}$ ), αλλά τις πραγματικές τιμές ( $\ln wage$ ). Παρατηρούμε ότι ο συντελεστής του μισθού έχει αλλάξει πλέον πρόσημο και είναι θετικός. Αν δηλαδή αφήσουμε εξ' αρχής εκτός ανάλυσης τους άνεργους (για τους οποίους δεν έχουμε παρατηρήσεις ούτε για τις ώρες εργασίας, ούτε για το μισθό) και δεν διορθώσουμε το πρόβλημα του selection bias (βάζοντας το συντελεστή mill's ratio και τον προβλεπόμενο μισθό), πλέον η ελαστικότητα προσφοράς εργασίας έχει θετικό πρόσημο, υπερσχύει δηλαδή το αποτέλεσμα υποκατάστασης.

**Πίνακας 11: OLS παλινδρόμηση (μόνο εργαζόμενοι και διόρθωση του selection bias)**

Indep. variables	OLS model	
	OLS 1 (Two-step Heckman prediction)	OLS 2 (MLE Heckman prediction)
lnwagehat	-13,80 (14,56)	-12,14 (14,83)
age	-1,49* (0,81)	-1,43* (0,81)
laborexp	0,77 (0,65)	0,73 (0,66)
gender	25,17*** (7,56)	23,83*** (8,14)
loweduc	8,19 (10,80)	11,88 (10,81)
mideduc	14,41 (9,45)	16,18* (9,54)
careerpersion	23,46*** (8,54)	22,70*** (8,76)
imr (lambda)	41,97* (24,65)	25,23 (25,17)
cons.	294,49*** (111,11)	281,10** (113,19)
R-squared	0,090	0,084
N	318	318

Notes: \*\*\*significant at 1%, \*\*at 5% (robust τυπικά σφάλματα σε παρένθεση)

**Πίνακας 12: OLS παλινδρόμηση (μόνο εργαζόμενοι και χωρίς διόρθωση του selection bias)**

Indep. variables	OLS
lnwage	11,45** (5,71)
age	-1,29* (0,77)
laborexp	0,72 (0,65)
gender	15,82** (6,84)
loweduc	22,25*** (8,62)
mideduc	22,43*** (8,24)
careerpersion	16,35* (8,33)
cons.	114,55** (45,24)
R-squared	0,098
N	318

Notes: \*\*\*significant at 1%, \*\*at 5%, at 10% (robust τυπικά σφάλματα σε παρένθεση)

Εν κατακλείδι, το πιο αξιόπιστο μοντέλο από τα παραπάνω είναι το λογοκριμένο μοντέλο tobit. Θα μπορούσε κάποιος να υπολογίσει εξ' αρχής ένα μοντέλο tobit, χωρίς να έχει εκτιμήσει προηγουμένως το μοντέλο Heckman, αλλά τότε θα υπήρχε πρόβλημα. Το πρόβλημα είναι ότι το tobit δεν ελέγχει για μεροληψία λόγω επιλογής δείγματος (*sample selection bias*), με αποτέλεσμα να μην είναι πολύ αξιόπιστα τα αποτελέσματα αν υπάρχει *selection bias*. Αν κάποιος έτρεχε εξ' αρχής tobit, ως ανεξάρτητη μεταβλητή του μισθού θα είχε το πραγματικό μισθό (*lnwage*). Αντίθετα, βάζοντας ως ανεξάρτητη τις προβλεπόμενες-εκτιμημένες τιμές του μισθού (*lnwagehat*), που εκτιμήθηκαν με βάση το μοντέλο Heckman, διορθώνεται το πρόβλημα του *selection bias*.



## Κεφάλαιο 6

### Συζήτηση-Σχολιασμός

Ο κλάδος των οικονομικών της εργασίας και ειδικότερα η εμπειρική εκτίμηση των ελαστικοτήτων προσφοράς εργασίας, αποτελεί ένα μεγάλο κομμάτι της οικονομικής επιστήμης, με αρκετές προκλήσεις. Το πιο συνηθισμένο πρόβλημα εμφανίζεται στη εκτίμηση της συνάρτησης προσφοράς εργασίας και αφορά τις μη παρατηρούμενες τιμές (*missing values*) στα δεδομένα κάποιων ομάδων ανθρώπων, όπως για παράδειγμα οι άνεργοι. Απόρροια αυτού του προβλήματος είναι και η μεροληψία επιλογής δείγματος (*sample selection bias*). Σε αυτές τις περιπτώσεις θα πρέπει να γίνεται πάντα έλεγχος για ύπαρξη ή μη της μεροληψίας, ειδικά όταν η αναλογία των μη παρατηρούμενων τιμών είναι μεγάλη σε σχέση με όλο το δείγμα.

Στη συγκεκριμένη μελέτη, το μέγεθος των μη παρατηρούμενων τιμών ήταν μικρό, αλλά σχετικά μικρό ήταν και το συνολικό μέγεθος του δείγματος. Γι' αυτό το λόγο κρίθηκε σκόπιμο να γίνει έλεγχος για πιθανή υιοθέτηση ενός μεροληπτικού δείγματος. Ο έλεγχος έγινε με το μοντέλο Heckman. Όπως είδαμε, υπάρχουν δύο τρόποι εκτίμησης του μοντέλου, με τον κάθε τρόπο να έχει τα υπέρ και τα κατά του. Η επικρατέστερη άποψη είναι ότι η εκτίμηση του μοντέλου Heckman με τη μέθοδο μεγίστης πιθανοφάνειας υπερτερεί έναντι της εκτίμησης των δύο σταδίων (Maddala, 1983; Davidson and MacKinnon, 1993). Μάλιστα, οι Davidson & MacKinnon (1993) υποστηρίζουν ότι το μοντέλο των δύο σταδίων θα πρέπει να χρησιμοποιείται μόνο για έλεγχο ύπαρξης της μεροληψίας, και αν υπάρχει, θα πρέπει να χρησιμοποιηθεί η μέθοδος μεγίστης πιθανοφάνειας για περαιτέρω ανάλυση. Στη παρούσα μελέτη βρέθηκε ότι υπάρχει πρόβλημα μεροληψίας, και όλη η υπόλοιπη ανάλυση βασίστηκε σε αυτή την υπόθεση.

Τα κύρια γνωρίσματα της εμπειρική ανάλυσης είναι α) ότι τα υποδείγματα εκτιμήθηκαν παράλληλα και με τις δύο μεθόδους εκτίμησης προκειμένου να φανούν οι διαφορές τους, και β) εκτιμήθηκαν εναλλακτικές εξισώσεις ελαχίστων τετραγώνων σαν να μην υπήρχε πρόβλημα μεροληψίας του δείγματος, προκειμένου να δούμε τις διαφορές τους με τα αντίστοιχα 'διορθωμένα' (ως προς τη μεροληψία) μοντέλα. Όσον αφορά το σημείο (α), αποδείχτηκε ότι και οι δύο μέθοδοι εκτίμησης του μοντέλου Heckman, στη συγκεκριμένη μελέτη λειτουργούν εξίσου καλά. Όσον αφορά το σημείο (β), ένα *ols* μοντέλο διορθωμένο ως προς τη μεροληψία δίνει αρνητική ελαστικότητα προσφοράς εργασίας, αλλά δεν είναι στατιστικά σημαντική. Αντίθετα, από το *ols* μοντέλο που δεν λαμβάνει υπόψη τη μεροληψία,

προκύπτει μία στατιστικά σημαντική ελαστικότητα, αλλά με θετικό πρόσημο. Το βασικό μοντέλο εκτίμησης της προσφοράς εργασίας είναι το tobit, και η ελαστικότητα που προκύπτει είναι στατιστικά σημαντική και έχει αρνητικό πρόσημο. Η εκτιμημένη ελαστικότητα των Ελλήνων της ώριμης ηλικίας (45-65 ετών) βρέθηκε ότι είναι -0,26. Χαρακτηριστικό επίσης είναι ότι το μορφωτικό επίπεδο, η ηλικία και η εμπειρία δεν είναι στατιστικά σημαντικές μεταβλητές, παρόλο που θεωρούνται βασικές μεταβλητές της εξίσωσης προσφοράς εργασίας και δεν λείπουν ποτέ από μία εκτίμηση, ή λείπει μόνο μία από τις τρεις. Στην εξίσωση μισθού του μοντέλου Heckman, η εκπαίδευση ήταν στατιστικά σημαντική, αλλά όχι και η ηλικία με την εμπειρία.

Ένας άλλος λόγος που στη συγκεκριμένη μελέτη υιοθετήθηκε η παραπάνω διαδικασία (πρώτα εκτίμηση ενός heckit μοντέλου και έπειτα η εκτίμηση ενός tobit μοντέλου) και ο οποίος λόγος δεν έχει δηλωθεί με σαφή τρόπο αλλά αφέθηκε να εννοηθεί 'σιωπηρά', είναι ότι με το μοντέλο heckit ελέγχονται ταυτόχρονα δύο προβλήματα. Εκτός από το πρόβλημα της μεροληψίας του επιλεγμένου δείγματος, ελέγχεται και το πρόβλημα πιθανής ενδογένειας μεταξύ των ωρών εργασίας και του μισθού. Σε περίπτωση δηλαδή που εκτιμούσε κάποιος εξ' αρχής το μοντέλο tobit με ανεξάρτητη μεταβλητή το πραγματικό μισθό, υπήρχε μεγάλη πιθανότητα να παρουσιαζόταν ενδογένεια μεταξύ των δύο μεταβλητών. Όμως η εκτίμηση του μοντέλου tobit με ανεξάρτητη μεταβλητή τις εκτιμημένες τιμές του μισθού και όχι τις πραγματικές τιμές, διορθώνει εκτός από το πρόβλημα της μεροληψίας και ένα πιθανό πρόβλημα ενδογένειας (Shaw, 1992; El-Hamidi, 2003).

## Κεφάλαιο 7

### Συμπεράσματα-Προτάσεις για περαιτέρω έρευνα

Η εκτιμημένη ελαστικότητα (-0,26) φαίνεται να συμφωνεί με τη βιβλιογραφία που υποστηρίζει ότι υπάρχουν ποικίλες εκτιμήσεις στη προσφορά εργασίας των αντρών (Botjas, 2003: σελ 76) και σε κάποιες περιπτώσεις μπορεί να είναι και αρνητικές (βλ. Πίνακα Β.2). Το 65% του δείγματός μας αποτελείται από άντρες, οπότε η εκτιμημένη αρνητική ελαστικότητα πιθανόν να οφείλεται στους άντρες. Επίσης η σχετική ανελαστικότητα είναι γνώριμο χαρακτηριστικό των αντρών αλλά και των γυναικών για τις χώρες της Ευρώπης (βλ. Πίνακα Α.1).

Το καλύτερο θα ήταν να εκτιμηθεί το μοντέλο προσφοράς εργασίας ξεχωριστά για τους άντρες και τις γυναίκες, προκειμένου να συγκρίνουμε τις αντίστοιχες ελαστικότητες με αυτές άλλων ερευνών από διάφορες χώρες, και αυτό γιατί η πλειοψηφία των μελετών εξετάζει τη προσφορά εργασίας μόνο των αντρών ή μόνο των γυναικών (κυρίως των γυναικών). Επομένως η ελαστικότητα προσφοράς εργασίας της συγκεκριμένης μελέτης δεν είναι άμεσα συγκρίσιμη με τις υπόλοιπες. Ο λόγος που δεν εκτιμήθηκε το μοντέλο προσφοράς εργασίας ξεχωριστά για τα δύο φύλα, είναι εξαιτίας των λίγων παρατηρήσεων.

Επίσης για περαιτέρω έρευνα, θα ήταν πολύ ενδιαφέρονσα μία μελέτη που εξετάζει τη διαχρονική συμπεριφορά των εργαζομένων, δηλαδή η χρήση ενός δυναμικού μοντέλου προσφοράς εργασίας. Βέβαια για να γίνει αυτό θα πρέπει να υπάρχουν διαθέσιμα panel δεδομένα. Τα δυναμικά μοντέλα είναι κατάλληλα και μπορούν να χρησιμοποιηθούν στην ανάλυση πολιτικής, ενώ τα στατικά μοντέλα είναι προτιμότερο να χρησιμοποιούνται για λόγους σύγκρισης. Παρόλα αυτά δεν σημαίνει ότι τα στατικά μοντέλα δεν μπορούν να χρησιμοποιηθούν σε αποφάσεις πολιτικής. Δηλαδή όταν διάφορες πολιτικές στην αγορά εργασίας οδηγούν σε μεταβολές του μισθού, τα ευρήματα του στατικού μοντέλου θα πρέπει να λαμβάνονται υπόψη, ειδικά αν στόχος είναι η αύξηση της απασχόλησης. Για παράδειγμα, τα ευρήματα της συγκεκριμένης μελέτης δείχνουν ότι σε μία αύξηση του μισθού, οι μεσήλικες εργαζόμενοι αντιδρούν μειώνοντας τις ώρες εργασίας τους. Επομένως μία πολιτική που οδηγεί σε αυξήσεις μισθών και παράλληλα έχει στόχο την αύξηση της απασχόλησης, ή τουλάχιστον τη διατήρησή της στα ίδια επίπεδα με πριν, δεν θα πρέπει να έχει εφαρμογή σε αυτή την ηλικιακή ομάδα (μέσης ηλικίας). Διαφορετικά, θα πρέπει να τροποποιηθεί η πολιτική για τα άτομα μέσης ηλικίας, έτσι ώστε να μην μειωθεί η απασχόλησή τους.

Συνήθως οι περισσότερες ελαστικότητες που εκτιμώνται στη βιβλιογραφία έχουν θετικό πρόσημο και οι αντίστοιχες ηλικιακές ομάδες που εξετάζονται είναι περίπου μεταξύ 20 και 65 ετών. Η αρνητική ελαστικότητα που βρέθηκε στη συγκεκριμένη ανάλυση για την Ελλάδα πιθανόν να οφείλεται στο ότι το υπό εξέταση δείγμα αφορούσε ανθρώπους που βρίσκονται στα τελευταία στάδια απασχόλησής τους (45-65 ετών). Είναι λογικό για κάποιον που έχει μεγάλη ηλικία να εργάζεται λιγότερες ώρες, παρόλο που μπορεί να αυξηθεί ο μισθός του. Ένας λόγος που οδηγεί σε μειωμένες ώρες εργασίας είναι ότι το άτομο πλησιάζει στη συνταξιοδότησή του. Έτσι αυτό που σκέφτεται είναι πότε θα σταματήσει να εργάζεται για να συνταξιοδοτηθεί. Ένας ακόμα λόγος είναι και τα προβλήματα υγείας. Τα άτομα μεγαλύτερης ηλικίας είναι πιθανότερο να έχουν σοβαρά προβλήματα υγείας από τους νέους. Όταν κάποιος έχει διάφορα προβλήματα υγείας, πιθανόν να μην μπορεί να εργαστεί πολλές ώρες. Εκτός αυτού, τα άτομα μεγαλύτερης ηλικίας κουράζονται πιο εύκολα από τους νεότερους, με αποτέλεσμα να εργάζονται λιγότερο απ' όσο ίσως να θέλουν. Η κούραση αυτή μπορεί να συνδέεται με κάποιο τωρινό πρόβλημα υγείας, αλλά μπορεί να οφείλεται και στη 'συσσωρευμένη' κούραση (ψυχολογική ή σωματική) από τα μέχρι τώρα χρόνια συνεχούς εργασίας. Οπότε βλέπουμε ότι υπάρχουν πολλοί λόγοι που μπορούν να εξηγήσουν το αρνητικό πρόσημο και τη σχετικά μικρή ελαστικότητα προσφοράς εργασίας της εμπειρικής μας μελέτης.

## ΕΛΛΗΝΙΚΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Χάλκος Γ. (2011). *Οικονομετρία: Θεωρία, Εφαρμογές & Χρήση Προγραμμάτων σε Η/Υ*, Εκδόσεις Gutenberg, Αθήνα.

Χαρίσης Κ.Ι. και Κιόχος Π.Α. (1997). *Θεωρία Δειγματοληψίας & Εφαρμογές*, Εκδόσεις Interbooks, Αθήνα.

## ΞΕΝΗ ΒΙΒΛΙΟΓΡΑΦΙΑ

Aaberge R., Colombino U. and Strom S. (1999). Labour Supply in Italy: An Empirical Analysis of Joint Household Decisions with Taxes and Quantity Constraints', *Journal of Applied Econometrics*, **14**, 403-422.

Bargain O., Orsini K. and Peichl A. (2011). Labor Supply Elasticities in Europe and the US, IZA Discussion Paper, No 5820.

Berg E.J. (1961). Backward-Sloping Labor Supply Functions in Dual Economies-the Africa Case, *Quarterly Journal of Economics*, **75**, 468-492.

Birch E.R. (2005). Studies of the Labour Supply of Australian Women: What Have We Learned?, *The Economic Record*, **81**, 65-84.

Bishop K., Heim B. and Mihaly K. (2009). Single Women's Labor Supply Elasticities: Trends and Policy Implications, *Industrial and Labor Relations Review*, **63**, 146-168.

Borjas G.J. (2003). *Οικονομικά της Εργασίας*, Εκδόσεις Κριτική, Αθήνα.

Browning M. and Chiappori P.A. (1998). Efficient Intra-Household Allocations: A General Characterization and Empirical Tests, *Econometrica*, **66**, 1241-1278.

Co C.Y., Gang I.N. and Yun M-S. (2005). Self-Employment and Wage Earning in Hungary, *Review of Development Economics*, **9**, 150-165.

Cogan J.F. (1981). Fixed Costs and Labor Supply, *Econometrica*, **49**, 945-963.

Daouli J., Demoussis M., Giannakopoulos N. (2004). Participation of Greek Married Women in Full-Time Paid Employment, *South Eastern Europe Journal of Economics*, **2**, 19-33.

Davidson R. and MacKinnon J.G. (1993). *Estimation and Inference in Econometrics*, Oxford University Press, Oxford.

del Campo C., Monteiro C.M.F and Soares J.O. (2008). The European Regional Policy and The Socio-Economic Diversity of European Regions: A multivariate Analysis, *European Journal of Operational Research*, **187**, 600-612.

Dessing M. (2002). Labor Supply, the Family and Poverty: the S-shaped Labor Supply Curve, *Journal of Economic Behavior & Organization*, **49**, 433-458.

Donni O. and Moreau N. (2007). Collective Labor Supply: A Single-Equation Model and Some Evidence from French Data, *Journal of Human Resources*, **42**, 214-246.

El-Hamidi F. (2003). Labor Supply of Egyptian Married Women when Self-Employment is an Option: Participation and Hours of Work, *Paper presented at the Annual Conference of the Middle East Economic Association (MEEA) and Allied Social Science Association (ASSA)*. January 3-5. Washington D.C.

Evers M., de Mooij R. and van Vuuren D. (2008). The Wage Elasticity of Labour Supply: A Synthesis of Empirical Estimates, *De Economist*, **156**, 25-43.

Gilley O.W. and Leone R.P. (1991). A Two-Stage Imputation Procedure for Item Nonresponse in Surveys, *Journal of Business Research*, **22**, 281-291.

Greene W.H. (1981). Sample Selection Bias as a Specification Error: Comment, *Econometrica*, **49**, 795-798.

Groshen E.L. and Williams D.R. (1992). White- and Blue-collar Jobs in the Recent Recession and Recovery: Who's Singing the Blues?, *Economic Review*, Federal Reserve Bank of Cleveland, **28**, 2-12.

Heckman J.J. (1976). The Common Structure of Statistical Models of Truncation, Sample Selection and Limited Dependent Variables and a Simple Estimator for Such Models, *The Annals of Economic and Social Measurement*, **5**, 475-492.

Heckman J.J. (1978). A Partial Survey of Recent Research on the Labor Supply of Women, *American Economic Review*, **68**, 200-207.

Heckman J.J. (1979). Sample Selection Bias as a Specification Error, *Econometrica*, **47**, 153-161.

Heckman J.J. (1993). What Has Been Learned About Labor Supply in the Past Twenty Years?, *American Economic Review*, **83**, 116-121.

Heckman J.J. and MaCurdy T.E. (1980). A Life Cycle Model of Female Labour Supply, *Review of Economic Studies*, **47**, 47-74.

Hill M.A. (1989). Female Labor Supply in Japan: Implications of the Informal Sector for Labor Force Participation and Hours of Work, *The Journal of Human Resources*, **24**, 143-161.

Jones A. (2007). *Applied Econometrics for Health Economists: A Practical Guide*, 2<sup>nd</sup> edition, Radcliffe Publishing, Oxford, Seattle.

Kauppinen-Toropainen K., Kandolin I. and Mutanen P. (1983). Job Dissatisfaction and Work-Related Exhaustion in Male and Female Work, *Journal of Occupational Behaviour*, **4**, 193-207.

Kidd M.P. and Ferko I. (2001). The Employment Effects of Gender Discrimination in Australia 1994-95, *The Economic Record*, **77**, 71-88.

Killingsworth M. (1983). *Labor Supply*, Cambridge University Press, New York.

Klevmarken N.A. (2005). Estimates of a Labour Supply Function Using Alternative Measures of Hours of Work, *European Economic Review*, **49**, 55-73.

Kuroda S. and Yamamoto I. (2008). Estimating Frisch labor supply elasticity in Japan, *Journal of The Japanese and International Economies*, **22**, 566-585.

LaRocco J.M., House J.S. and French Jr. J.R.P. (1980). Social Support, Occupational Stress, and Health, *Journal of Health and Social Behavior*, **21**, 202-218.

Lee L-F. (1982). Some Approaches to the Correction of Selectivity Bias, *The Review of Economic Studies*, **49**, 355-372.

Leibowitz A. and Klerman J.A. (1995). Explaining Changes in Married Mothers' Employment over Time, *Demography*, **32**, 365-378.

Leung S.F. and Yu S. (1996). On the Choice Between Sample Selection and Two-Part Models, *Journal of Econometrics*, **72**, 197-229.

Leung S.F. and Yu S. (2000). Collinearity and Two-Step Estimation of Sample Selection Models: Problems, Origins, and Remedies, *Computational Economics*, **15**, 173-199.

MaCurdy T., Green D. and Paarsch H. (1990). Assessing Empirical Approaches for Analyzing Taxes and Labor Supply, *Journal of Human Resources*, **25**, 415-490.

Maddala G.S. (1983). *Limited-Dependent and Qualitative Variables in Econometrics*, Cambridge University Press, New York, NY.

Maglad N.A. (1998). Female Labour Supply in Sudan, *The African Economic Research Consortium (AERC)*, Special Paper 30, The Regal Press, Nairobi, Kenya.

Melino A. (1982). Testing for Sample Selection Bias, *The Review of Economic Studies*, **49**, 151-153.

Mizala A., Romaguera P. and Henriquez P. (1999). Female Labor Supply in Chile, Working paper 58, Centro de Economia Aplicada, Universidad de Chile.

Mroz T.A. (1987). The Sensitivity of An Empirical Model of Married Women's Hours of Work to Economic and Statistical Assumptions, *Econometrica*, **55**, 765-799.

Nam S. (1991). Determinants of Female Labor Force Participation: A Study of Seoul, South Korea, 1970-1980, *Sociological Forum*, **6**, 641-659.

Nawata K. (1993). A note on the Estimation of Models with Sample-Selection Biases, *Economics Letters*, **42**, 15-24.

Nawata K. (1994). Estimation of Sample Selection Bias Models by the Maximum Likelihood Estimator and Heckman's two-step Estimator, *Economics Letters*, **45**, 33-40.

Nawata K. and Ii M. (2004). Estimation of the Labor Participation and Wage Equation Model of Japanese Married Women by the Simultaneous Maximum Likelihood Method, *Journal of the Japanese and International Economies*, **18**, 301-315.

Nawata K. and McAleer M. (2001). Size Characteristics of Tests for Sample Selection Bias: a Monte Carlo Comparison and Empirical Example, *Econometric Reviews*, **20**, 105-112.



- Nawata K. and Nagase N. (1996). Estimation of Sample Selection Bias models, *Econometric Reviews*, **15**, 387-400.
- Nelson F.D. (1984). Efficiency of the Two-Step Estimator for Models with Endogenous Sample Selection, *Journal of Econometrics*, **24**, 181-196.
- Olsen R.J. (1980). A Least Squares Correction for Selectivity Bias, *Econometrica*, **48**, 1815-1820.
- Olsen R.J. (1982). Distributional Tests for Selectivity Bias and a More Robust Likelihood Estimator, *International Economic Review*, **23**, 223-240.
- Paarsch H.J. (1984). A Monte Carlo Comparison of Estimators for Censored Regression Models, *Journal of Econometrics*, **24**, 197-213.
- Parker S.C., Belghitar Y. and Barmby T. (2005). Wage Uncertainty and the Labour Supply of Self-Employed Workers, *The Economic Journal*, **115**, 190-207.
- Powell J.L. (1984). Least Absolute Deviations Estimation for the Censored Regression Model, *Journal of Econometrics*, **25**, 303-325.
- Puhani P.A. (2000). The Heckman Correction for Sample Selection and its Critique, *Journal of Economic Surveys*, **14**, 53-68.
- Sartori A. (2003). An Estimator for Some Binary-Outcome Selection Models Without Exclusion Restrictions, *Political Analysis*, **11**, 111-138.
- Schaffner J.A. (2002). Heteroskedastic Sample Selection and Developing-Country Wage Equations, *The Review of Economics and Statistics*, **84**, 269-280.
- Schultz T.P. (1980). Estimate Labor Supply Functions for Married Women. In: Female Labor Supply: Theory and Estimation, Smith JP (Eds), Princeton University Press, Princeton, 1980. pp. 25-89.
- Sharif M. (1991). Poverty and the Forward-Falling Labor Supply Function: A Microeconomic Analysis, *World Development*, **19**, 1075-1093.
- Shaw K.L. (1992). The Life-Cycle Labor Supply of Married Women and its Implications for Household Income Inequality, *Economic Inquiry*, **XXX**, 659-672.

Takahashi S., Kawade M. and Kato R.R. (2009). Labor Supply of Japanese Married Women: Sensitivity Analysis and a New Estimate, CRR working papers, Shiga University, Japan, No. A-17

Wetzels C. (2005). Supply and Price of Childcare and Female Labour Force Participation in the Netherlands, *Labour (special issue)*, **19**, 171-209.

Winship C. and Mare R.D. (1992). Models for Sample Selection Bias, *Annual Review of Sociology*, **18**, 327-350.

Yamagata T. and Orme C.D. (2005). On Testing Sample Selection Bias Under the Multicollinearity Problem, *Econometric Reviews*, **24**, 467-481.

## **Διευθύνσεις στο διαδίκτυο**

<http://www.uis.unesco.org/Library/Documents/isc97-en.pdf>

## ΠΑΡΑΡΤΗΜΑ Α

Πίνακας Α.1: Ονομασίες μεταβλητών

Όνομα μεταβλητής	Ερμηνεία μεταβλητής
hours	Ώρες εργασίας (μηνιαίες)
lnwage	Φυσικός λογάριθμος του μισθού (μηνιαίος)
age	Ηλικία (σε έτη)
gender	Φύλο (1: άνδρας, 0: γυναίκα)
laborexp	Εργασιακή εμπειρία (σε έτη)
loweduc	Επίπεδο εκπαίδευσης (1: χαμηλή εκπαίδευση, 0: οτιδήποτε άλλο)
mideduc	Επίπεδο εκπαίδευσης (1: μεσαία εκπαίδευση, 0: οτιδήποτε άλλο)
higheduc	Επίπεδο εκπαίδευσης (1: υψηλή εκπαίδευση, 0: οτιδήποτε άλλο) (μένει εκτός ανάλυσης)
housemembers	Συνολικός αριθμός ατόμων που υπάρχουν στο νοικοκυριό
privsector	Ιδιωτικός τομέας (1: αν εργάζεται στον ιδιωτικό τομέα, 0: αν εργάζεται στο δημόσιο τομέα)
whitecollar	Φύση της εργασίας (1: αν εργάζεται σε μη χειρωνακτικές θέσεις απασχόλησης, 0: αν εργάζεται σε χειρωνακτικές θέσεις απασχόλησης)
personscontrib	Συνολικός αριθμός ατόμων που συνεισφέρουν στο εισόδημα
careerperson	Άτομο καριέρας (1: αν ο ερωτώμενος θεωρεί ότι είναι άτομο καριέρας, 0: αν δεν θεωρεί ότι είναι άτομο καριέρας)

**Πίνακας Α.2: Βοηθητική οls παλινδρόμηση για έλεγχο πολυσυγγραμμικότητας**

Indep. variables	inverse mill's ratio (two-step)	inverse mill's ratio (ML)
age	0,0024 (0,0016)	-0,000 (0,002)
gender	-0,237*** (0,014)	-0,245*** (0,015)
laborexp	0,0037*** (0,0012)	0,007*** (0,002)
loweduc	0,08*** (0,017)	0,069*** (0,018)
mideduc	0,02 (0,017)	0,023 (0,017)
housemembers	0,051*** (0,005)	0,06*** (0,005)
privsector	0,178*** (0,014)	0,166*** (0,015)
whitecollar	-0,107*** (0,015)	-0,062*** (0,016)
personscontrib	-0,119*** (0,009)	-0,109*** (0,01)
cons.	0,042 (0,073)	0,033 (0,077)
R-squared	0,728	0,693
adj. R-squared	0,721	0,685
N	358	358

Notes: \*\*\*significant at 1%, (τυπικά σφάλματα σε παρένθεση)

**Πίνακας Α.3: Συντελεστές συσχέτισης**

	Inwagehat	age	laborexp	gender	loweduc	mideduc	careerperson
Inwagehat	1						
age	-0,016	1					
laborexp	0,028	0,691	1				
gender	0,293	0,098	0,205	1			
loweduc	-0,327	0,089	0,278	0,023	1		
mideduc	-0,084	-0,099	-0,044	-0,038	-0,524	1	
careerperson	0,194	0,006	-0,106	0,141	-0,262	-0,064	1

## ΠΑΡΑΡΤΗΜΑ Β

**Πίνακας Β.1: Ελαστικότητες προσφοράς εργασίας στην Ευρώπη**

Χώρα	Μελέτη	Δεδομένα	Ελαστικότητες μισθού	
			Άντρες	Γυναίκες
Γερμανία	Steiner & Wrohlich (2004)	SOEP (2003)		[0.20, 0.36]
	Haan & Steiner (2005)	SOEP (2002)	[0.08, 0.31]	[0.02, 0.24]
	Clauss & Schnabel (2006)	SOEP (2004/5)	0.23	0.38
	Bargain et al. (2009)	SOEP (2003)	[0.10, 0.20]	[0.06, 0.16]
	Fuest et al. (2008)	SOEP (2004)	0.28	0.28
Ιταλία	Aaberge et al. (2002)	SHIW (1993)	0.11	0.10
Ολλανδία	Euwals & van Soest (1999)	SOEP (1988)	[0.03, 0.18]	[0.03, 0.45]

Πηγή: Bargain et al. (2011)

**Πίνακας Β.2: Ελαστικότητες προσφοράς εργασίας στις ΗΠΑ**

Μελέτη	Δεδομένα	Στοιχεία	Ελαστικότητες μισθού	
			Άντρες	Γυναίκες
Cogan (1981)	National Longitudinal Study of Mature Women (1967)	Παντρεμένες γυναίκες 30-35 ετών		[0.86, 2.40]
Hausman (1981)	PSID 1975	Παντρεμένες γυναίκες		[0.90, 1.00]
Triest (1990)	PSID 1983	Παντρεμένες γυναίκες 25-55 ετών		[0.03, 0.28]
MaCurdy et al. (1990)	PSID 1975	Παντρεμένοι άντρες 25-55 ετών	[-0.24, 0.03]	
Pencavel (2002)	CPS 1999	Παντρεμένοι και ελεύθεροι άντρες	[0.12, 0.25]	
Devereux (2003)	Census and PSID	Όλοι οι άντρες	[-0.022, 0.017]	
Devereux (2004)	PUMS 1980/1990	Παντρεμένα ζευγάρια	[0.00, 0.07]	[0.17, 0.38]
Blau & Kahn (2007)	CPS 1980	Παντρεμένοι άντρες και γυναίκες 25-54 ετών	[0.01, 0.07]	[0.77, 0.88]
	CPS 1990		[0.10, 0.14]	[0.58, 0.64]
	CPS 2000		[0.04, 0.10]	[0.36, 0.41]
Heim (2009)	PSID 2001	Ζευγάρια	[0.07, 0.18]	[0.24, 0.33]

Πηγή: Bargain et al. (2011)