

ΣΕΜΙΝΑΡΙΑ ΕΚΠΟΝΗΣΗΣ ΔΙΠΛΩΜΑΤΙΚΗΣ ΕΡΓΑΣΙΑΣ - RESEARCH METHODS

Διδάσκων: Αναπλ. καθ. Λογιστικής Ορέστης Βλησμάς

Εφαρμογή 6^η

Δεδομένα πάνελ και ανάλυση παλινδρόμησης στο **STATA**

Ακολουθήστε τα παρακάτω βήματα:

Βήμα 1°:

1. Μεταβείτε στο e-class του σεμιναρίου. Στο folder με ονομασία Laboratory του e-class αναζητήστε το αρχείο με ονομασία Application_6.dta. Αποθηκεύεται το τοπικά στον υπολογιστή σας και ανοίξτε το στο πρόγραμμα λογισμικού **STATA**. **ΕΠΙΣΗΜΑΝΣΗ:** Το αρχείο Application_6.dta είναι το τελικό αρχείο δεδομένων **STATA** όπως έχει προκύψει μετά την Εφαρμογή 4 με αναφορά την ερευνητική εργασία του Roychowdhury (2006).
2. Μεταβείτε στο e-class του σεμιναρίου. Στο folder με ονομασία Laboratory του e-class αναζητήστε το αρχείο με ονομασία Code_Application_5.txt. Αποθηκεύεται το τοπικά στον υπολογιστή σας και ανοίξτε το στο πρόγραμμα λογισμικού STATA.
3. Μεταβείτε στο e-class του σεμιναρίου. Στο folder με ονομασία Library_Stata αναζητήστε και ανοίξτε το αρχείο με ονομασία Xristikos_Odigos_Stata. Αποθηκεύεται το τοπικά στον υπολογιστή σας. Το αρχείο αυτό θα το έχετε ανοικτό στην διάρκεια της εκτέλεσης της εφαρμογής προκειμένου να το συμβουλευέστε αν χρειαστεί.

Βήμα 2°: Αναγγελία δεδομένων πάνελ

Δηλώστε στο **STATA** ότι τα δεδομένα σας είναι τύπου πάνελ με την ακόλουθη εντολή:
xtset id year

ΠΑΡΑΤΗΡΗΣΕΙΣ

1. Όταν ανοίγεται ένα αρχείο dta το οποίο εμπεριέχει δεδομένα τύπου πάνελ (και φυσικά η δομή των δεδομένων είναι τύπου Long) ΘΑ ΠΡΕΠΕΙ ΟΠΩΣΔΗΠΟΤΕ να ΤΟ δηλώσετε στο STATA. Αυτό γίνεται με τη σύνταξη της εντολής: xtset ΔΗΛΩΤΙΚΗ ΜΕΤΑΒΛΗΤΗ ΠΕΡΙΠΤΩΣΗΣ ΔΗΛΩΤΙΚΗ ΜΕΤΑΒΛΗΤΗ ΧΡΟΝΟΥ. Στην περίπτωσή μας η μεταβλητή id εκτελεί τον ρόλο της δηλωτικής μεταβλητής περίπτωσης και η μεταβλητή year τη δηλωτική μεταβλητή χρόνου.

Βήμα 3°: Δημιουργία πρόσθετων μεταβλητών βασικού μοντέλου γραμμικής παλινδρόμησης της ερευνητική εργασία του Roychowdhury (2006)

Με βάση αναφοράς την ερευνητική εργασία του Roychowdhury (2006) θα πρέπει να δημιουργηθούν κάποιες πρόσθετες μεταβλητές για την εκτίμηση του βασικού μοντέλου γραμμικής παλινδρόμησης. Για το λόγο αυτό αρχικά πρέπει να μελετήσουμε το σχετικό κείμενο της ερευνητικής εργασίας:

5.1. Comparison of suspect firm-years with the rest of the sample

If firm-years that report profits just above zero undertake activities that adversely affect their CFO, then the abnormal CFO for these firm-years, calculated using the industry-year model described in Section 4.2, should be negative compared to the rest of the sample. To test this, I estimate the following regression:

$$Y_t = \alpha + \beta_1 (SIZE)_{t-1} + \beta_2 (MTB)_{t-1} + \beta_3 (Net\ income)_t + \beta_4 (SUSPECT_NI)_t + \varepsilon_t. \quad (6)$$

In this case, the dependent variable, Y_t is abnormal CFO in period t . Regression (6) is also estimated with abnormal production costs and abnormal discretionary expenses as the dependent variables. $SUSPECT_NI$ is an indicator variable that is set equal to one if firm-years belong to the earnings category just right of zero, and zero otherwise.

To control for systematic variation in abnormal CFO, production costs and discretionary expenses with growth opportunities and size, the regression includes two control variables: MTB and $SIZE$. MTB , or the market-to-book ratio, is the ratio of market value of equity to book value of equity. $SIZE$ is the logarithm of the market value of equity at the beginning of the year. Dechow et al. (1995, 1996) argue that abnormal accruals calculated using conventional, non-discretionary-accruals models have measurement error positively correlated with firm performance. To address the possibility that abnormal values from my estimation models have measurement error correlated with performance, I include net income as a control variable in the regressions.²² The net income figure is scaled by lagged total assets, so it is similar to return-on-assets (ROA).²³ Since the dependent variables are essentially deviations from 'normal' levels within an industry-year, all the control variables in the regressions are also expressed as deviations from the respective industry-year means.

Από την ανάγνωση του κειμένου διαπιστώνουμε ότι πρέπει να δημιουργηθούν:

1. Μία μεταβλητή που θα σηματοδοτεί τα «ύποπτα» έτη (την οποία θα την ονομάσουμε **suspect**).
2. Μία μεταβλητή μεγέθους της εταιρείας ορισμένη ως ο λογάριθμος της χρηματιστηριακής αξίας των ιδίων κεφαλαίων της εταιρείας της προηγούμενης χρήσης μείον το αντίστοιχο μέσο όρο του κλάδου δραστηριότητας (την οποία θα την ονομάσουμε **l.sizek**).
3. Μία μεταβλητή μεγέθους της εταιρείας ορισμένη ως ο λόγος της χρηματιστηριακής αξίας των ιδίων κεφαλαίων της εταιρείας προς την αντίστοιχη λογιστική της προηγούμενης χρήσης μείον το αντίστοιχο μέσο όρο του κλάδου δραστηριότητας (την οποία θα την ονομάσουμε **l.mtbk**).
4. Μία μεταβλητή οικονομικής απόδοσης της εταιρείας ορισμένη ως των ιδίων κεφαλαίων της εταιρείας η αποδοτικότητα των ιδίων κεφαλαίων μείον το αντίστοιχο μέσο όρο του κλάδου δραστηριότητας (την οποία θα την ονομάσουμε **nik**).

Για το λόγο αυτό εκτελέστε στο STATA τα ακόλουθα:

```
generate size=log 10(mv)
generate mtb=mv/tas
by id: gen lsize=L.size
by id:gen mtbl=L.mtb
egen sizemean=mean(size), by (industry)
egen nimean=mean(ni), by (industry)
egen mtbmean=mean(mtb), by (industry)
generate sizek=size-sizemean
generate nik=ni-nimean
```

generate mtbk=mtb-mtbmean

Βήμα 4°: Αρχική εκτίμηση γραμμικής παλινδρόμησης

Η βασική παλινδρόμηση θα εκτελεσθεί αρχικά λαμβάνοντας υπόψη ότι έχουν δημιουργηθεί οι ακόλουθες τρεις μεταβλητές (δες Εφαρμογή 4) για την χειραγώγηση λογιστικών κερδών μέσω λειτουργικών αποφάσεων (real earnings management):

1. **abnormalcfor**: ασυνήθιστα χαμηλές καθαρές ταμειακές ροές από λειτουργικές δραστηριότητες εξαιτίας της μόχλευσης των πωλήσεων.
2. **abnormalDiscrr**: υπερβάλλουσα μείωση των δαπανών που υπόκεινται στη διακριτική ευχέρεια των διοικητικών στελεχών.
3. **abnormalproducr**: υπερβάλλουσα μείωση του ανά μονάδα κόστους εξαιτίας της αδικαιολόγητης αύξησης της παραγωγής

Ειδικότερα εκτελέστε τις ακόλουθες εντολές:

```
xtreg abnormalcfor l.sizek l.mtbk nik suspect  
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect  
xtreg abnormalproducr l.sizek l.mtbk nik suspect
```

Κάντε μία σύντομη αξιολόγηση για κάθε ένα από τα εκτιμημένα γραμμικά μοντέλα παλινδρόμησης.

Βήμα 5°: Έλεγχος παραδοχών γραμμικού μοντέλου παλινδρόμησης

Στο βήμα αυτό θα πραγματοποιηθούν οι ενέργειες ελέγχου παραδοχών του γραμμικού μοντέλου παλινδρόμησης.

Ενέργεια 1 – Έλεγχος κανονικής κατανομής καταλοίπων με swilk, sfrancia, sktest

Για να ελεγχθεί η υπόθεση της κανονικής κατανομής των καταλοίπων εκτελούνται οι έλεγχοι swilk, sfrancia, sktest με τις ακόλουθες εντολές:

```
xtreg abnormalcfor l.sizek l.mtbk nik suspect  
predict e, residuals  
swilk e  
sfrancia e  
sktest e  
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect  
estimates store e  
swilk e  
sfrancia e  
sktest e  
xtreg abnormalproducr l.sizek l.mtbk nik suspect  
estimates store e  
swilk e  
sfrancia e  
sktest e
```

Αν το p-value είναι χαμηλότερο από 0,05 τότε απορρίπτεται η υπόθεση της κανονικής κατανομής.

Ενέργεια 2 – Έλεγχος πολυσυγγραμμικότητας με vif

Για να ελεγχθεί η ύπαρξη της πολυσυγγραμμικότητας εκτελείται ο έλεγχος vif ως εξής:

```
reg abnormalcfor l.sizek l.mtbk nik suspect  
vif  
reg abnormalDiscrr l.sizek l.mtbk nik suspect  
vif  
reg abnormalproducr l.sizek l.mtbk nik suspect  
vif
```

Ενέργεια 3 – Έλεγχος ετεροσκεδαστικότητας με hettest και imtest, whit.

Για το hettest εκτελέστε τις ακόλουθες εντολές:

```
reg abnormalcfor l.sizek l.mtbk nik suspect  
estat hettest  
reg abnormalDiscrr l.sizek l.mtbk nik suspect  
estat hettest  
reg abnormalproducr l.sizek l.mtbk nik suspect  
estat hettest
```

Για το imtest, white εκτελέστε τις ακόλουθες εντολές:

```
reg abnormalcfor l.sizek l.mtbk nik suspect  
estat imtest, white  
reg abnormalDiscrr l.sizek l.mtbk nik suspect  
estat imtest, white  
reg abnormalproducr l.sizek l.mtbk nik suspect  
estat imtest, whit
```

Αν το p-value είναι χαμηλότερο από 0,05 τότε υπάρχει ετεροσκεδαστικότητα. Σε περίπτωση που διαπιστωθεί ετεροσκεδαστικότητα η προσαρμογή γίνεται με την επιλογή robust και η σχετική σύνταξη έχει ως εξής:

```
xtreg abnormalcfor l.sizek l.mtbk nik suspect, robust  
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect, robust  
xtreg abnormalproducr l.sizek l.mtbk nik suspect, robust
```

Ενέργεια 4 – Έλεγχος αυτοσυσχέτισης με bgodfrey

Για το bgodfrey εκτελέστε τις ακόλουθες εντολές:

```
generate time = _n  
tsset time (η εντολή αυτή δηλώνει χρονολογική σειρά)  
reg abnormalcfor l.sizek l.mtbk nik suspect  
estat bgodfrey, lag(1)  
reg abnormalDiscrr l.sizek l.mtbk nik suspect  
estat bgodfrey, lag(1)  
reg abnormalproducr l.sizek l.mtbk nik suspect  
estat bgodfrey, lag(1)
```

Αν το p-value είναι χαμηλότερο από 0,05 τότε υπάρχει αυτοσυσχέτιση. Σε περίπτωση που διαπιστωθεί αυτοσυσχέτιση η προσαρμογή γίνεται με την επιλογή cluster(id) όπου id είναι η μεταβλητή δηλωτική της εταιρείας και η σχετική σύνταξη έχει ως εξής:

```
xtreg abnormalcfor l.sizek l.mtbk nik suspect, cluster(id)
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect, cluster(id)
xtreg abnormalproducr l.sizek l.mtbk nik suspect, cluster(id)
```

Ενέργεια 5 – Έλεγχος σταθερών έναντι τυχαίων επιδράσεων με hausman

Για το hausman εκτελέστε τις ακόλουθες εντολές:

```
xtreg abnormalcfor l.sizek l.mtbk nik suspect, fe
estimates store fixed
xtreg abnormalcfor l.sizek l.mtbk nik suspect, re
estimates store random
hausman fixed random
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect, fe
estimates store fixed
xtreg abnormalDiscrr l.sizek l.mtbk nik suspect, re
estimates store random
hausman fixed random
xtreg abnormalproducr l.sizek l.mtbk nik suspect, fe
estimates store fixed
xtreg abnormalproducr l.sizek l.mtbk nik suspect, re
estimates store random
hausman fixed random
```

Αν το p-value είναι χαμηλότερο από 0,05 τότε επιλέξτε fe (σταθερές επιδράσεις).

Ενέργεια 6 – Οριστικοποίηση επιλογών παλινδρόμησης xtreg

Οι επιλογές fe, re, robust, cluster(Firm_id) είναι δυνατόν να χρησιμοποιηθούν συνδυασμένα ανάλογα με τη περίπτωση και το πρόβλημα που διαπιστώσαμε στις προηγούμενες ενέργειες. Έστω ότι επιθυμούμε να εκτιμήσουμε ένα υπόδειγμα (πολυμεταβλητής) παλινδρόμησης που αφορά μια εξαρτημένη μεταβλητή (vrY) και ένα πλήθος ανεξάρτητων μεταβλητών (έστω τρεις με την ονομασία vrX1, vrX2 και vrX3 αντίστοιχα) με δεδομένα πάνελ. Ο παρακάτω πίνακας δίδει διάφορες επιλογές κατά την εκτέλεση της εντολής xtreg.

Επιδράσεις	Ετεροσκεδαστικότητα	Αυτοσυσχέτιση	Σύνταξη
Σταθερές	Όχι	Όχι	xtreg vrY vrX 1 vrX2 vrX3, fe
Σταθερές	Ναι	Όχι	xtreg vrY vrX1 vrX2 vrX3, fe robust
Σταθερές	Όχι	Ναι	xtreg vrY vrX1 vrX2 vrX3, fe cluster(FirmId)
Σταθερές	Ναι	Ναι	xtreg vrY vrX1 vrX2 vrX3, fe robust cluster(FirmId)
Τυχαίες	Όχι	Όχι	xtreg vrY vrX1 vrX2 vrX3, re
Τυχαίες	Ναι	Όχι	xtreg vrY vrX1 vrX2 vrX3, re robust
Τυχαίες	Όχι	Ναι	xtreg vrY vrX1 vrX2 vrX3, re cluster(FirmId)
Τυχαίες	Ναι	Ναι	xtreg vrY vrX1 vrX2 vrX3, re robust cluster(FirmId)