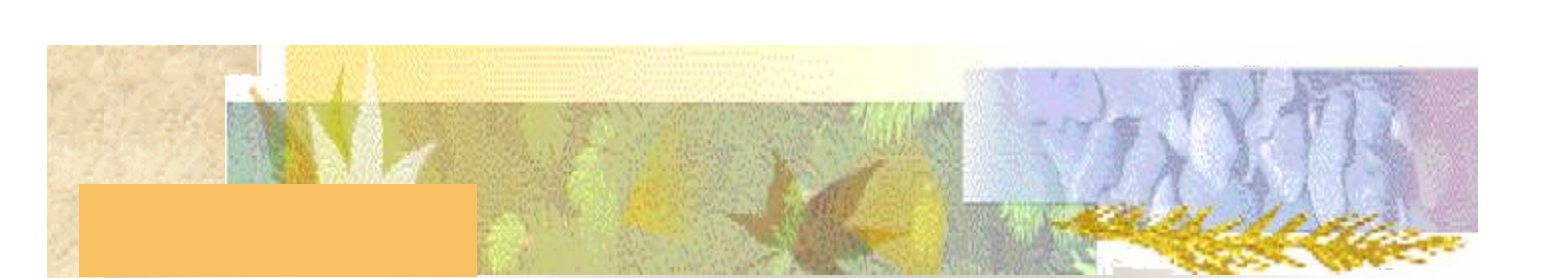


**ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΤΜΗΜΑ ΤΡΑΠΕΖΙΚΗΣ ΚΑΙ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΔΙΟΙΚΗΤΙΚΗΣ**

Εκπόνηση διπλωματικής εργασίας

«Μοντέλα Αξιολόγησης Πιστωτικού Κινδύνου»

ΙΟΥΝΙΟΣ 2002



ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΕΙΡΑΙΩΣ
ΜΕΤΑΠΤΥΧΙΑΚΟ ΤΜΗΜΑ ΤΡΑΠΕΖΙΚΗΣ ΚΑΙ
ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΗΣ ΔΙΟΙΚΗΤΙΚΗΣ

- Επιβλέπων καθηγητής: Μπένος Αλέξανδρος
- Μεταπτυχιακή φοιτήτρια: Κασμά Έλενα

ΙΟΥΝΙΟΣ 2002

I. ΕΙΣΑΓΩΓΗ

1. ΔΑΝΕΙΟΔΟΤΗΣΗ.....	σελ. 1
2. ΘΕΜΑ ΕΡΓΑΣΙΑΣ.....	σελ. 1

II. ΜΟΝΤΕΛΑ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ ΘΕΩΡΗΤΙΚΗ ΠΡΟΣΕΓΓΙΣΗ

1. ΠΙΣΤΩΤΙΚΟΣ ΚΙΝΔΥΝΟΣ

1.1. ΟΡΙΣΜΟΣ ΚΑΙ ΠΡΟΕΛΕΥΣΗ.....	σελ. 3
1.2. ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ΣΤΟ ΠΑΡΕΛΘΟΝ.....	σελ. 3
1.3. ΕΠΕΚΤΑΣΗ ΚΑΙ ΔΙΑΔΟΣΗ ΠΙΣΤΩΣΗΣ.....	σελ. 4
1.4. ΑΙΤΙΕΣ ΑΛΛΑΓΗΣ ΚΑΙ ΕΞΕΛΙΞΗΣ ΣΤΗΝ ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ.....	σελ. 4
- ΑΛΛΑΓΗ ΚΟΥΛΤΟΥΡΑΣ	
- ΑΛΛΑΓΗ ΠΟΛΙΤΙΚΩΝ ΚΑΙ ΟΙΚΟΝΟΜΙΚΩΝ ΣΥΝΘΗΚΩΝ	
- ΠΑΡΑΔΕΙΓΜΑΤΑ ΑΘΕΤΗΣΗΣ	
1.5. ΕΞΕΛΙΞΕΙΣ ΣΤΗ ΔΙΑΧΕΙΡΙΣΗ ΤΟΥ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ.....	σελ. 9

2. ΠΙΣΤΩΤΙΚΗ ΑΝΑΛΥΣΗ

2.1. ΔΙΑΚΡΙΣΗ ΜΕΤΑΞΥ ΠΟΙΟΤΙΚΗΣ ΚΑΙ ΠΟΣΟΤΙΚΗΣ.....	σελ. 11
2.2. ΠΟΣΟΤΙΚΗ ΠΙΣΤΩΤΙΚΗ ΑΝΑΛΥΣΗ.....	σελ. 11
- ΜΕΤΑΣΤΡΟΦΗ ΤΗΣ ΑΝΑΛΥΣΗΣ ΑΠΟ ΤΟΥΣ ΙΣΟΛΟΓΙΣΜΟΥΣ ΣΤΙΣ ΚΑΤΑΣΤΑΣΕΙΣ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ	
- ΑΝΑΛΥΣΗ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ ΜΕ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΥΣ ΔΕΙΚΤΕΣ	
- ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΑΝΑΛΥΣΗΣ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ ΜΕ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΥΣ ΔΕΙΚΤΕΣ	
2.3. ΣΤΡΟΦΗ ΑΠΟ ΧΡΗΣΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΩΝ ΔΕΙΚΤΩΝ ΣΕ ΧΡΗΣΗ ΟΙΚΟΝΟΜΕΤΡΙΚΩΝ ΜΟΝΤΕΛΩΝ (FROM UNIVARIATE TO MULTIVARIATE)	σελ. 16

3. ΜΟΝΤΕΛΑ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

3.1. ΧΡΗΣΙΜΟΤΗΤΑ ΜΟΝΤΕΛΩΝ.....	σελ. 17
3.2. ΚΑΤΑΣΚΕΥΗ ΜΟΝΤΕΛΩΝ.....	σελ. 17
3.3. ΤΕΧΝΙΚΕΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ.....	σελ. 17
3.4. ΕΦΑΡΜΟΓΕΣ ΚΑΙ ΠΛΕΟΝΕΚΤΗΜΑΤΑ.....	σελ. 18
3.5. ΟΡΙΣΜΟΙ ΜΟΝΤΕΛΩΝ.....	σελ. 19
α) DISCRIMINANT ANALYSIS.....	σελ. 19
β) LINEAR PROBABILITY MODEL	σελ. 20
γ) PROBIT & LOGIT MODEL.....	σελ. 22

3.6.ΣΥΝΟΠΤΙΚΗ ΑΝΑΦΟΡΑ ΣΤΙΣ ΕΜΠΕΙΡΙΚΕΣ ΜΕΛΕΤΕΣ ΤΟΥ ΠΑΡΕΛΘΟΝΤΟΣ.....	σελ.24
3.7.ΑΝΑΛΥΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΠΡΟΓΕΝΕΣΤΕΡΩΝ ΕΜΠΕΙΡΙΚΩΝ ΜΕΛΕΤΩΝ....	σελ.26
a) Univariate approach / Beaver.....	σελ.26
b) MDA	σελ.26
i.Altman Z-score 1968.....	σελ.26
ii.Zeta model 1977.....	σελ.29
iii.Altman Z-score / Bank of England 1982.....	σελ.31
iv.Joint discriminant functions and neural networks 1994.....	σελ.32
c) LPM.....	σελ.34
i.Linear Probability Model and Discriminant Analysis- εφαρμογή στην Ελλάδα 1984	σελ.34
d) Logit.....	σελ.36
i.Logit Regression approach / Martin 1977.....	σελ.36
e) Newer models	σελ.38
3.8.ΙΔΙΑΙΤΕΡΟΤΗΤΕΣ ΜΟΝΤΕΛΩΝ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ.....	σελ.38
3.8.1.ΕΜΠΕΙΡΙΚΑ ΕΝΑΝΤΙ ΘΕΩΡΗΤΙΚΩΝ ΜΟΝΤΕΛΩΝ.....	σελ.38
3.8.2. BIAS	σελ.40

III. ΕΜΠΕΙΡΙΚΟΣ ΕΛΕΓΧΟΣ

1. ΣΤΟΧΟΣ ΕΡΓΑΣΙΑΣ.....	σελ.42
2. ΟΡΙΣΜΟΣ ΠΡΟΒΛΗΜΑΤΟΣ.....	σελ.42
3. ΑΝΤΙΣΤΟΙΧΙΣΗ-ΔΗΜΙΟΥΡΓΙΑ ΔΕΙΓΜΑΤΩΝ.....	σελ.42
4. ΔΗΜΙΟΥΡΓΙΑ ΜΟΝΤΕΛΩΝ.....	σελ.44
5. ΕΜΠΕΙΡΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ.....	σελ.48
5.1. Probit	σελ.48
<u>5.1.1.Στατιστική σημαντικότητα μεταβλητών</u>	σελ.48
<u>5.1.2.Συσχέτιση μεταβλητών</u>	σελ.50
<u>5.1.3.Στατιστική σημαντικότητα μοντέλου- Τεστ</u>	σελ.53
<i>Αρχικό δείγμα (1yr prior)</i>	
<u>5.1.4.Ακρίβεια πρόβλεψης μοντέλου</u>	σελ.54
<u>5.1.5.Goodness of fit test</u>	σελ.56
<i>Τεστ 1: Αρχικό δείγμα 2yrs prior</i>	
<i>Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior</i>	
<i>Τεστ 3: 2^ο δείγμα 2yrs prior</i>	
<i>Τεστ 4: 2^ο δείγμα 1yr prior</i>	
<i>Τεστ 5: Αρχικό δείγμα (1yr prior)- ΕΠΑΝΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ</i>	
<u>5.1.6.Γενικά συμπεράσματα για το probit</u>	σελ.66

5.2. Logit	σελ.67
<u>5.2.1. Στατιστική σημαντικότητα μεταβλητών</u>	σελ.67
<u>5.2.2. Συσχέτιση μεταβλητών</u>	σελ.68
<u>5.2.3. Στατιστική σημαντικότητα μοντέλου- Τεστ</u>	σελ.70
<i>Αρχικό δείγμα (1yr prior)</i>	
<u>5.2.4. Ακρίβεια πρόβλεψης μοντέλου</u>	σελ.70
<u>5.2.5. Goodness of fit test</u>	σελ.72
<i>Τεστ 1: Αρχικό δείγμα 2yrs prior</i>	
<i>Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior</i>	
<i>Τεστ 3: 2^ο δείγμα 2yrs prior</i>	
<i>Τεστ 4: 2^ο δείγμα 1yr prior</i>	
<i>Τεστ 5: Αρχικό δείγμα (1yr prior)- ΕΠΑΝΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ</i>	
<u>5.2.6. Γενικά συμπεράσματα για probit και logit</u>	σελ.82
5.3. MDA	σελ.84
<u>5.3.1. Υποθέσεις μοντέλου</u>	σελ.84
<u>5.3.2. Στατιστική σημαντικότητα μεταβλητών</u>	σελ.85
<i>-Συσχέτιση μεταβλητών</i>	
<i>-Stepwise statistics</i>	
<u>5.3.3. Στατιστική σημαντικότητα μοντέλου- Τεστ</u>	σελ.97
<i>Τεστ 1: Αρχικό δείγμα 2yrs prior</i>	
<i>Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior</i>	
<i>Τεστ 3: 2^ο δείγμα 2yrs prior</i>	
<i>Τεστ 4: 2^ο δείγμα 1yr prior</i>	
<u>5.3.4. Γενικά συμπεράσματα για MDA</u>	σελ.111
<u>5.3.5. Γενικά συμπεράσματα για όλα τα μοντέλα</u>	σελ.111

IV. ΠΑΡΑΡΤΗΜΑ

ΠΙΝΑΚΑΣ 1

ΠΙΝΑΚΑΣ 2

Ι.ΕΙΣΑΓΩΓΗ

ΔΑΝΕΙΟΔΟΤΗΣΗ

Η δανειοδότηση ή αλλιώς πίστωση εμφανίστηκε από τα αρχαία κιάλας χρόνια ως μια πρωταρχική οικονομική ανάγκη της ανθρώπινης κοινωνίας και λειτούργησε αναμφισβήτητα εως σήμερα ως κινητήριος μοχλός για την τεχνολογική προώθηση και την οικονομική ανάπτυξη του ανθρώπινου γένους.

Γενικά, μια διαδικασία δανειοδότησης θα μπορούσαμε να πούμε ότι αναπτύσσεται σε τέσσερα στάδια: την αίτηση για χορήγηση δανείου από την ενδιαφερόμενη εταιρία, την ανάλυση της οικονομικής κατάστασης και κατ' επέκταση της πιστοληπτικής ικανότητας του πελάτη, την τιμολόγηση και δόμηση του δανείου σύμφωνα με την παραπάνω ανάλυση και τέλος την αποπληρωμή του δανείου. Η πιστωτική ανάλυση που αποτελεί το δεύτερο στάδιο της διαδικασίας, αποσκοπεί στον περιορισμό του αριθμού των επισφαλών δανείων στο ελάχιστο δυνατό και στην έγκαιρη ανίχνευση οποιουδήποτε προβλήματος στην εξόφληση του δανείου, ώστε ο δανειστής να έχει τη δυνατότητα να ζητήσει νωρίτερα την εξόφληση ή και να αποσυρθεί από την διαδικασία.

Μια σύγχρονη πιστωτική ανάλυση αναπτύσσεται σε δύο βασικά επίπεδα : την ποιοτική ανάλυση της δανειζόμενης επιχείρησης και την ποσοτική ανάλυση των οικονομικών της καταστάσεων και των ταμειακών της ροών. Σκοπός της πρώτης είναι η ανίχνευση και κατανόηση της συγκριτικής θέσης της επιχείρησης μέσα στην αγορά, των ανταγωνιστικών πιέσεων που δέχεται, της δυσκολίας εισόδου στον κλάδο της, του επιπέδου τεχνολογικής της εξέλιξης, της νοοτροπίας και κουλτούρας που την διέπουν, των προσόντων και ικανοτήτων των βασικών στελεχών της. Σκοπός της δεύτερης από την άλλη είναι η απόσπαση από τις οικονομικές της καταστάσεις σημαντικών πληροφοριών κυρίως για τις μεταβολές των πραγματικών ταμειακών ροών καθώς και η εξαγωγή ειδικών χρηματοοικονομικών δεικτών.

ΘΕΜΑ ΕΡΓΑΣΙΑΣ

Σε αυτήν την εργασία σκοπεύουμε να ασχοληθούμε με το δεύτερο επίπεδο της πιστωτικής ανάλυσης, που αφορά την ανάλυση ποσοτικών οικονομικών στοιχείων.

Προτού φτάσουμε στο πρακτικό και ερευνητικό μέρος της εργασίας μας, θα αναφερθούμε στη σταδιακή εξέλιξη της πίστωσης και του πιστωτικού κινδύνου από τη στιγμή της γέννησης τους μέχρι σήμερα. Θα επισημάνουμε την υπέρμετρη διάδοση της πίστωσης που οφείλεται στην αλλαγή των κοινωνικών αντιλήψεων και την κοινωνική αποδοχή της και στην επακόλουθη αύξηση του πιστωτικού κινδύνου. Στη συνέχεια, με την παράθεση και κατάλληλων παραδειγμάτων από την οικονομική πραγματικότητα, θα εξηγήσουμε πώς οι παραπάνω κοινωνικές και οικονομικές εξελίξεις έκαναν ανεπαρκή την μέχρι τότε υφιστάμενη πιστωτική πολιτική. Η αναγκαιότητα για μια στροφή από την ποιοτική και υποκειμενική ανάλυση της πιστοληπτικής ικανότητας στην ποσοτική ανάλυση οικονομικών δεδομένων θα τονιστεί ιδιαίτερα. Τέλος, αφού θα έχουμε επικεντρώσει πλέον την προσοχή μας στην ποσοτική πιστωτική ανάλυση, θα προσπαθήσουμε να εξηγήσουμε την επιπλέον στροφή από την ποσοτική ανάλυση χρηματοοικονομικών δεικτών (μονομεταβλητή μέθοδος) στην χρήση οικονομετρικών μοντέλων (πολυμεταβλητές μέθοδοι), για να καταλήξουμε έτσι στο κύριο θέμα της εργασίας μας.

Αυτό που θα μας απασχολήσει κυρίως, είναι η εφαρμογή και συγκριτική μελέτη συγκεκριμένων οικονομετρικών μοντέλων μέτρησης πιστωτικού κινδύνου και η εξαγωγή συμπερασμάτων για την απόλυτη και σχετική αποτελεσματικότητά τους στα ελληνικά δεδομένα. Μετά λοιπόν από την θεωρητική προσέγγιση του πιστωτικού κινδύνου και των διαφόρων τεχνικών που χρησιμοποιήθηκαν για την αντιμετώπισή του, θα επικεντρωθούμε στην πιο πρόσφατη από αυτές που είναι τα οικονομετρικά μοντέλα και θα προχωρήσουμε σε μια θεωρητική ανάλυση των τελευταίων και των μέχρι τώρα εφαρμογών τους, δίνοντας έμφαση κυρίως σε αυτά που θα χρησιμοποιήσουμε.

Με την τελευταία αυτή αναφορά τελειώνει και το θεωρητικό κομμάτι της εργασίας μας. Όσον αφορά το ερευνητικό τώρα μέρος, σκοπεύουμε να εφαρμόσουμε τρία οικονομετρικά μοντέλα σε δεδομένα ελληνικών εταιριών. Τα μοντέλα που θα χρησιμοποιήσουμε είναι: η Multiple Discriminant Analysis (MDA), το Probit model και το Logit model. Τα μοντέλα αυτά θα εφαρμοστούν σε στοιχεία 42 χρεωκοπημένων και 42 υγιών εταιριών αντίστοιχου κλάδου και μεγέθους. Βασιζόμενοι στις εμπειρικές μελέτες προηγούμενων αναλυτών, θα ακολουθήσουμε τα ίδια με αυτούς βήματα σε κάθε διαδικασία, ελπίζοντας να καταλήξουμε σε εύλογα συμπεράσματα για την ικανότητα αυτών των τριών μοντέλων να προβλέπουν σωστά τον πιστωτικό κίνδυνο μέσα στην ελληνική επιχειρησιακή αγορά.

1. ΠΙΣΤΩΤΙΚΟΣ ΚΙΝΔΥΝΟΣ

1.1. ΟΡΙΣΜΟΣ ΚΑΙ ΠΡΟΕΛΕΥΣΗ

Ο πιστωτικός κίνδυνος αποτελεί την αρχαιότερη μορφή κινδύνου στα πλαίσια των οικονομικών αγορών. Αν η πίστωση μπορεί να οριστεί ως η προσδοκία ενός χρηματικού ποσού μέσα σε ένα ορισμένο χρονικό διάστημα, ο πιστωτικός κίνδυνος είναι η πιθανότητα αυτή ακριβώς η προσδοκία να μην επαληθευτεί.

Ο πιστωτικός κίνδυνος είναι τόσο παλιός όσο ο ίδιος ο δανεισμός. Η προέλευση του χρονολογείται τουλάχιστον γύρω στα 1800 π.Χ. και παραμένει ουσιαστικά अपαράλλακτος από την εποχή των αρχαίων Αιγυπτίων: σήμερα, όπως και τότε, διακρίνεται πάντα από ένα στοιχείο αβεβαιότητας για το κατά πόσον ένας δεδομένος δανειστής θα αποπληρώσει ένα συγκεκριμένο δάνειο.

1.2. ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ΣΤΟ ΠΑΡΕΛΘΟΝ

Από την εποχή της πρώιμης δημιουργίας οργανωμένων τραπεζών, 700 χρόνια πριν, οι τράπεζες υπήρξαν τα πρωταρχικά πιστωτικά ιδρύματα της κοινωνίας. Η διαχείριση του πιστωτικού κινδύνου αποτέλεσε από την αρχή τον βασικό πυρήνα των δραστηριοτήτων τους.

Από τις απαρχές της σύγχρονης τραπεζικής στην μεσαιωνική Ευρώπη, το ιδιαίτερο χαρακτηριστικό των τραπεζών υπήρξε η λειτουργία τους ως ιδρύματα που δέχονταν καταθέσεις. Μέχρι και σήμερα, οι περισσότεροι τραπεζικοί κανονισμοί εξακολουθούν να προσδιορίζουν την έννοια της τράπεζας με βάση αυτή τη λειτουργία. Έχοντας λοιπόν συμφωνήσει να ενεργούν ως καταθετήρια ιδρύματα, άρχισαν να συλλογίζονται τρόπους με τους οποίους θα μπορούσαν να χρησιμοποιήσουν τις καταθέσεις μέχρις ότου οι πελάτες γυρίσουν να τις αποσύρουν. Ο πρώτος και ιδιαίτερα επικερδής από αυτούς ήταν ο δανεισμός. Οι τραπεζίτες δεν άργησαν να αντιληφθούν πως δανείζοντας τις προσόδους από τις καταθέσεις των πελατών τους θα μπορούσαν να δημιουργήσουν μία παράλληλη και κερδοφόρα δουλειά. Κι ενώ μέχρι τότε συνήθιζαν να επιβαρύνουν τους καταθέτες τους με ένα αντίτιμο για την υπηρεσία φύλαξης των χρημάτων τους, άρχισαν στη συνέχεια να τους αμοιβούν προκειμένου να προσελκύσουν τις καταθέσεις τους και να τις αξιοποιήσουν μέσω της διαδικασίας του δανεισμού. Κατ'αυτόν τον τρόπο, πλήρωναν ένα αντίτιμο,

γνωστό ως τόκο, στους καταθέτες τους, χρέωναν ένα υψηλότερο στους δανειζόμενους και κέρδιζαν από τη διαφορά των δύο. Η εναλλακτική αυτή δραστηριότητα όμως του δανεισμού που υιοθετήθηκε πολύ νωρίς από τις τράπεζες, περιείχε και έναν αναπόφευκτο κίνδυνο: να μην λάβουν ποτέ πίσω μερικά από τα κεφάλαια που δάνειζαν. Για πολλά χρόνια, το επιτόκιο καταθέσεων δεν προσδιοριζόταν από τις δυνάμεις της αγοράς, περιορίζοντας έτσι τον πιστωτικό κίνδυνο σε καθορισμένα στενά όρια.

Παραδοσιακά, οι τραπεζίτες και οι υπόλοιποι δανειστές αντιμετώπισαν την εκτίμηση του πιστωτικού κινδύνου κατά τον ίδιο περίπου τρόπο που ένας ράφτης προσεγγίζει την δημιουργία ενός ενδύματος στα ακριβή μέτρα του πελάτη: μετρώντας προσεκτικά τις ανάγκες και ικανότητες του πελάτη ώστε να βεβαιωθεί ότι η χρηματοδότηση «είναι στα μέτρα του». Η σημερινή τακτική δεν διαφέρει κατ' ουσία τουλάχιστον από αυτήν που χρησιμοποιούσαν οι αρχαιότερες τράπεζες.

Επειδή ακριβώς είναι ευκολότερο να σχεδιάσεις ένα κοστούμι για έναν πελάτη που ήδη γνωρίζεις, οι τράπεζες στο παρελθόν, έδωσαν μεγαλύτερη έμφαση στις πελατειακές τους σχέσεις (relationship banking). Τους ενδιέφερε περισσότερο η σχέση τους με τον πελάτη παρά η κερδοφορία του συγκεκριμένου δανείου ή οι συνέπειες αυτής της συναλλαγής στο συνολικό χαρτοφυλάκιο δανείων της τράπεζας. Τις τελευταίες δεκαετίες όμως, η παραδοσιακή αυτή τακτική οδήγησε σε τραγικά και απαράδεκτα αποτελέσματα: οι τράπεζες σημείωσαν παταγώδεις αποτυχίες στην προσπάθειά τους να διαχειριστούν τον πιστωτικό κίνδυνο, τη στιγμή μάλιστα που η διάδοσή του αυξανόταν δραματικά.

1.3. ΕΠΕΚΤΑΣΗ ΚΑΙ ΔΙΑΔΟΣΗ ΠΙΣΤΩΣΗΣ

Ο πιστωτικός κίνδυνος τα τελευταία χρόνια εμφανίζεται, ως αναπόδραστο στοιχείο της τραπεζικής δραστηριότητας, όλο και με μεγαλύτερη συχνότητα, εξαιτίας της υπέρμετρης θα λέγαμε διάδοσης της πίστωσης. Ειδικά στην Αμερική, εταιρίες δανείζονται για να προβούν σε αγορές και να μεγεθυνθούν, μικρότερες επιχειρήσεις δανείζονται για να επεκτείνουν τη δραστηριότητά τους και χιλιάδες άτομα χρησιμοποιούν πίστωση για να αγοράσουν σπίτια, αυτοκίνητα, ρούχα, ακόμη και φαγητό. Το στοιχείο του πιστωτικού κινδύνου εμφανίζεται από τη στιγμή που ένα άτομο λαμβάνει ένα προϊόν ή μια υπηρεσία δίχως να προβαίνει σε άμεση πληρωμή για αυτό. Οι τηλεφωνικές εταιρίες και οι εταιρίες παροχής ηλεκτρικού ρεύματος ή ύδρευσης αναλαμβάνουν διαρκώς πιστωτικό κίνδυνο από τους συνδρομητές τους. Το ίδιο συμβαίνει και με τους εκδότες των πιστωτικών καρτών ή τους δανειστές στην περίπτωση δανείων με υποθήκη. Όσον αφορά τον επιχειρησιακό τομέα, όλες ανεξαιρέτως οι εταιρίες προσπαθούν να πουλήσουν τα προϊόντα ή τις υπηρεσίες τους υπό κάποιους όρους πίστωσης. Κάθε φορά που συμβαίνει κάτι τέτοιο, αποδέχονται και έναν πιστωτικό κίνδυνο, ο οποίος μπορεί να διαρκέσει από μερικές ώρες έως και πολλά χρόνια.

1.4.ΑΙΤΙΕΣ ΑΛΛΑΓΗΣ ΚΑΙ ΕΞΕΛΙΞΗΣ ΣΤΗΝ ΑΝΤΙΜΕΤΩΠΙΣΗ ΤΟΥ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

-ΑΛΛΑΓΗ ΚΟΥΛΤΟΥΡΑΣ

Η ‘έκρηξη’ αυτή του φαινομένου της πίστωσης και κατά συνέπεια της ανάληψης του πιστωτικού κινδύνου, συνοδεύτηκε και επιταχύνθηκε από μία δραματική αλλαγή στις κοινωνικές αντιλήψεις. Στα παλαιότερα χρόνια, ο δανειζόμενος αντιμετωπιζόταν με ντροπή και καταφρόνηση. Σήμερα, οποιοδήποτε είδος ντροπής συνυφασμένο με τον δανεισμό έχει εξαφανιστεί. Ακόμη και το λεξιλόγιο που χρησιμοποιούμε πλέον για να περιγράψουμε την πίστωση αντανακλά μια μεγάλη αλλαγή νοοτροπίας. Η λέξη χρεώστης υποδηλώνει ακόμη μίζερια και ντροπή. Η λέξη δανειζόμενος επίσης. Σήμερα όμως δεν χρειάζεται να αντιμετωπίσουμε τους εαυτούς μας ως χρεώστες ή δανειστές : μπορούμε απλά να θεωρούμαστε ως άτομα που χρησιμοποιούν μόχλευση, μία λέξη που παραπέμπει σε άλλες έννοιες καθόλου ευκαταφρόνητες. Η μόχλευση υποδηλώνει ότι κατέχουμε αρκετή ευφυΐα και ικανότητα για να αξιοποιήσουμε ένα εργαλείο που πολλαπλασιάζει τη δύναμη μας. Η χρήση μόχλευσης αποτελεί ένα κατόρθωμα για το οποίο έχουμε το δικαίωμα να περηφανευόμαστε και όχι μια πράξη ανάγκης που μας ατιμάζει.

Όλοι οι πολίτες των ανεπτυγμένων αλλά και αναπτυσσόμενων χωρών, βομβαρδίζονται στην κυριολεξία από αμέτρητες κατεύθυνσεις με προσκλήσεις για αύξηση των πιστώσεών τους. Οι αυτοκινητοβιομηχανίες προσελκύουν τους πελάτες τους με χαμηλών επιτοκίων και χαλαρών όρων δάνεια για αγορά αυτοκινήτου, οι εκδότες πιστωτικών καρτών κατακλύζουν τα πραγματικά και ηλεκτρονικά ταχυδρομεία με ολοένα και πιο ανταγωνιστικές προσφορές, οι λιανοπωλητές προκαλούν τους καταναλωτές να ανοίξουν πιστωτικούς λογαριασμούς με την προσφορά εκπτώσεων στις πρώτες αγορές τους. Τελικά, αν υφίσταται ακόμη κάποιο ίχνος ντροπής στις σημερινές συναλλαγές, αυτό συνδέεται πλέον μάλλον με τα μετρητά. Ο έμπορος που θα επιμείνει να πληρώσεις με μετρητά μπορεί κάλλιστα να θίγει την φερεγγυότητα και ακεραιότητά σου.

Η αλλαγή αυτή στη νοοτροπία της αγοράς αντανακλάται βεβαίως και στην επιχειρησιακή σφαίρα. Υψηλά και καλοπληρωμένα στελέχη προσλαμβάνονται καθημερινά για την ανίχνευση και εύρεση χρηματικών πόρων με σκοπό τη μόχλευση της εταιρίας τους.

Ακόμη και η χρεωκοπία δεν ακούγεται πλέον τόσο υποτιμητική. Μία προς αποφυγήν ντροπιαστική και μοιραία κατάληξη του παρελθόντος, γίνεται σήμερα ευρέως αποδεκτή ως μια λογική στρατηγική επιλογή.

Η κοινωνική αποδοχή και η εκτεταμένη διαθεσιμότητα της πίστωσης διευκολύνει τη ζωή της σύγχρονης κοινωνίας. Η πίστωση δίνει τη δυνατότητα σε άτομα ακόμη και χαμηλής οικονομικής στάθμης να αποκτήσουν αυτοκίνητα, σπίτια, καταναλωτικά αγαθά, κι αυτό με τη σειρά του μειώνει την ανεργία και προωθεί την συνολική οικονομική ανάπτυξη. Η πίστωση δίνει την δυνατότητα στις επιχειρήσεις να μεγεθυνθούν και να αναπτυχθούν. Επιτρέπει στις κυβερνήσεις, τις πόλεις και τους φορείς τους να καλύψουν τις ανάγκες του κοινωνικού συνόλου.

Και παρά την σημερινή ελαστικότερη στάση απέναντι στη χρεωκοπία, οι κίνδυνοι των πιστωτών τις περισσότερες φορές δεν είναι ιδιαίτερα υψηλοί. Οι καταναλωτές γνωρίζουν ότι αν δεν πληρώσουν τους λογαριασμούς τηλεφώνου ή ηλεκτρικού ρεύματος, η σύνδεσή τους θα κοπεί. Οι ιδιοκτήτες σπιτιών ξέρουν πως αν σταματήσουν να πληρώνουν τις δόσεις των στεγαστικών δανείων τους θα χάσουν τα σπίτια τους. Οι επιχειρήσεις επίσης γνωρίζουν ότι αν δεν πληρώσουν τους προμηθευτές τους θα πάνουν να λαμβάνουν πλέον προμήθειες. Τα άτομα και οι εταιρίες αντιλαμβάνονται ότι η πρόσβασή τους στην πίστωση εξαρτάται από την επιτυχία και απόδοσή τους στο να αναταποκριθούν στις τρέχουσες υποχρεώσεις τους. Από τη στιγμή λοιπόν που η πίστωση είναι ζωτική για αυτούς θα κάνουν το παν για να ξεπληρώνουν τα χρέη τους.

Παρ'όλ'αυτα βέβαια, κάθε επιχείρηση που δέχεται πίστωση στις καθημερινές της εργασίες υφίσταται κάποιες απώλειες. Οι περισσότερες εταιρίες υπολογίζουν το συνήθες επίπεδο απωλειών τους και προσπαθούν να διαμορφώσουν την πιστωτική πολιτική τους έτσι ώστε να μην υπερβαίνουν αυτό το επίπεδο. Απώλειες τέτοιου είδους και μεγέθους, δεν αποτελούν απειλή.

Όταν η οικονομία είναι ανθηρή, η επιρροή της πίστωσης εμφανίζεται μικρή. Σε οικονομικές υφέσεις όμως, μπορεί να μεταβληθεί σε σοβαρό πρόβλημα και να οδηγήσει σε απώλειες θέσεων και ανεργία, σε καθυστερήσεις πληρωμών δανειακών δόσεων και δόσεων πιστωτικών καρτών, μέχρι και σε χρεωκοπίες ολοκληρων εταιριών. Αν μάλιστα η ύφεση είναι αρκετά βαθιά και διαρκής, ακόμη και τα πιστωτικά ιδρύματα αρχίζουν να απειλούνται. Όσο πιο εκτεταμένο το χρέος τόσο πιο μεγάλος ο κίνδυνος.

-ΑΛΛΑΓΗ ΠΟΛΙΤΙΚΩΝ ΚΑΙ ΟΙΚΟΝΟΜΙΚΩΝ ΣΥΝΘΗΚΩΝ

Ο πιστωτικός κίνδυνος αυξήθηκε κατά εκθετικό θα λέγαμε τρόπο τη δεκαετία του 1990 παράλληλα με τις δραματικές οικονομικές, πολιτικές και τεχνολογικές αλλαγές που σημειώθηκαν σε ολόκληρο τον κόσμο. Στις προηγούμενες δεκαετίες, τα γεωπολιτικά σύνορα και οι κυβερνητικοί κανονισμοί περιόριζαν την κίνηση των κεφαλαίων. Από τη στιγμή όμως που, στις αρχές του 1970, οι συναλλαγματικές ισοτιμίες έγιναν κυμαινόμενες, οι χρηματοοικονομικές αγορές άρχισαν να γνωρίζουν μια σταθερά αυξανόμενη αποκανονικοποίηση και απελευθέρωση. Το γεγονός αυτό οδήγησε σε έναν αυξανόμενο ανταγωνισμό ανάμεσα στα πιστωτικά ιδρύματα και σε μια διαδοχική κατάργηση των συνόρων που χώριζαν τις τράπεζες, τις ασφαλιστικές εταιρίες και τις εταιρίες αμοιβαίων κεφαλαίων.

Αλλά και στο γεωπολιτικό επίπεδο σημειώθηκαν δραματικές αλλαγές. Η διπολική ένταση μεταξύ κομμουνισμού και καπιταλισμού εξασθένησε και μαζί με αυτήν η ανάγκη για υπέρογκους στρατιωτικούς προϋπολογισμούς. Η έμφαση σήμερα δίνεται στην οικονομική ανάπτυξη και τη βελτίωση του βιοτικού επιπέδου.

Η ρευστότητα ανά τον κόσμο αυξήθηκε, κάτι που ερμηνεύεται ως αύξηση των πηγών, των χρήσεων και των χρηστών κεφαλαίου. Οι

συναλλαγές που λαμβάνουν χώρα σήμερα παρουσιάζουν ποικιλία μοναδική. Σε πολλές περιπτώσεις πια, ο πιστωτικός κίνδυνος πρέπει να αναλυθεί για συμβαλλόμενους που διαθέτουν από λίγη έως καθόλου πιστωτική ιστορία. Ο πιστωτικός κίνδυνος με άλλα λόγια έγινε πιο πολύπλοκος, ενώ ταυτόχρονα η δυνατότητα των κυβερνήσεων για εγγυήσεις και οικονομικές στηρίξεις έχει μειωθεί.

-ΠΑΡΑΔΕΙΓΜΑΤΑ ΑΘΕΤΗΣΗΣ

Τα αποτελέσματα των παραπάνω μεταβολών διαφαίνονται καθαρά από τα πλήθος παραδείγματα αθετήσεως υποχρεώσεων που σημειώθηκαν παγκοσμίως τις δεκαετίες του 1970 και 1980. Οι Ηνωμένες Πολιτείες συγκεκριμένα, παρουσίασαν ρεκόρ αθετήσεων σε τραπεζικά δάνεια και εταιρικές ομολογίες στα μέσα της δεκαετίας του 1980. Τα ποσοστά υπήρξαν τόσο μεγάλα που οι αναλυτές της εποχής υποστήριζαν ότι οι junk bonds και οι τραπεζικές αγορές δανείων πολύ πιθανόν να αφανίζονταν.

Διαθέσιμα στοιχεία υποδηλώνουν ότι για πολλά χρόνια, οι αμερικανικές τράπεζες υποτιμούσαν συστηματικά τον πιστωτικό κίνδυνο. Οι λόγοι για την ανεπαρκή αυτή αποτίμηση είναι ποικίλοι: ορισμένες τράπεζες χρησιμοποίησαν τα εμπορικά δάνεια ως στοιχεία προώθησης άλλων προϊόντων τους περισσότερο επικερδών. Επιπλέον, οι τράπεζες υστερούσαν σε συστήματα μέτρησης της κερδοφορίας των πελατειακών σχέσεων σε όρους απόδοσης κεφαλαίου. Μία ακόμη αιτία μπορεί απλά να είναι η άγνοια: ακόμη και σήμερα, οι τράπεζες παρουσιάζουν σημαντικές ελλείψεις σε στοιχεία που αφορούν την πιστωτική και δανειοδοτική εμπειρία του παρελθόντος. Σε πολλές περιπτώσεις υποχρεώνονται να δανειστούν τέτοιου είδους στοιχεία από τον ιδιωτικό τομέα και τις αγορές ομολόγων, προκειμένου να κατανοήσουν την δουλειά τους και να τιμολογήσουν σωστά τα δάνειά τους.

Γεγονότα των πρόσφατων δεκαετιών έδειξαν ότι η κρίση των Αμερικανών δανειστών κάθε άλλο παρά αλάνθαστη υπήρξε. Οι Αμερικανικές τράπεζες διέπραξαν σοβαρά λάθη όσον αφορά τον πιστωτικό τομέα. Παρόλο που ποικίλοι παράγοντες συνετέλεσαν στη δημιουργία των πρόσφατων τραπεζικών κρίσεων, η ανεπαρκής πιστωτική πολιτική υπήρξε σίγουρα ένας από τους σημαντικότερους. Δανείζοντας τις χώρες της Λατινικής Αμερικής τη δεκαετία του 1970 και τις εταιρίες διαχείρισης ακινήτων τη δεκαετία του 1980, παρασυρόμενες από τον ελκυστικό λόγο κινδύνου/απόδοσης που πρόσφεραν οι συγκεκριμένες επενδύσεις και αγνοώντας τα πιστωτικά όρια που οι ίδιες είχαν θέσει, στήριξαν τις αποφάσεις τους στην παραδοσιακή πιστωτική

μεθοδολογία: εκτίμησαν εξατομικευμένους κινδύνους και επικεντρώθηκαν στην πίστωση πελατών με τους οποίους είχαν μακροχρόνιες σχέσεις. Οι τεχνικές αυτές όμως οδήγησαν σε παταγώδεις αποτυχίες. Τόσο στην περίπτωση της Λατινικής Αμερικής όσο και στην περίπτωση των εταιριών διαχείρισης ακινήτων, οι τράπεζες ζημιώθηκαν επειδή επέλεξαν όχι λάθος ατομικούς κινδύνους, αλλά λάθος τομείς. Οι κρίσεις βέβαια και στις δύο περιπτώσεις, προήλθαν και από έναν αριθμό εξωτερικών παραγόντων. Στη συγκεκριμένη περίπτωση για παράδειγμα της Λατινικής Αμερικής, πολλές από τις δυσκολίες που αντιμετώπισαν αυτές οι χώρες, οφείλονταν σε εξωτερικά σοκ, όπως ήταν η απότομη πτώση παγκοσμίως των τιμών των προϊόντων και τα υψηλά επιτόκια που συνόδευσαν την περιοριστική νομισματική πολιτική των Ηνωμένων Πολιτειών σε μια προσπάθεια να ελεγχθεί ο πληθωρισμός. Στην περίπτωση των εταιριών διαχείρισης ακινήτων από την άλλη, οι τράπεζες προέβησαν στον δανεισμό τεράστιων χρηματικών κεφαλαίων στηριζόμενες σε αισιόδοξες προβλέψεις για τη μελλοντική εξέλιξη των τιμών των ακινήτων και την ανάληψη μεγάλου μέρους των δανείων αυτών από ασφαλιστικές εταιρίες, προσδοκίες όμως που δεν επαληθεύτηκαν. Οι παραπάνω εξωτερικοί παράγοντες, δεν μπορούν παρολαυτά να απαλλάξουν τις αμερικανικές τράπεζες από τις δικές τους ευθύνες. Δεν υπάρχει πραγματικά εύκολος τρόπος για να εξηγήσουμε το γεγονός ότι μία ομάδα ευφυών επαγγελματιών με μεγάλα κεφάλαια διαθέσιμα είχε πειστεί ότι οι χώρες της Λατινικής Αμερικής θα μπορούσαν να απαλλαγούν κάποια στιγμή από τα υπέρογκα ποσά χρέους που τις βάραιναν.

Κατά ένα μέρος λοιπόν, οι δύο παραπάνω κρίσεις αντανakλούν μία αποτυχία των τραπεζών να καθιερώσουν και να διατηρήσουν μια αποτελεσματική πιστωτική κουλτούρα.

Αλλά και στην Ευρώπη, αποτυχημένες δανειοδοτήσεις προκάλεσαν μια τραπεζική κρίση ανάλογη αυτής των Ηνωμένων Πολιτειών. Εξαιτίας μεγάλων πιστωτικών ζημιών, μεγάλες τράπεζες στη Γαλλία και την Ισπανία έφτασαν στα πρόθυρα της χρεωκοπίας. Προβλήματα πιστωτικού κινδύνου σημειώθηκαν και στα τραπεζικά συστήματα της Νορβηγίας, της Σουηδίας και της Φινλανδίας. Όσον αφορά τον υπόλοιπο πλανήτη, παρόμοιες δυσκολίες μικρότερου ή μεγαλύτερου μεγέθους, παρουσιάστηκαν και στις οικονομίες της Κορέας, της Ταϊλάνδης, της Μαλαισίας και της Ιαπωνίας.

Όλες οι παραπάνω μεταβολές αναφέρθηκαν για να εξηγήσουμε τις αιτίες που προκάλεσαν τη διαχρονική εξέλιξη της πίστωσης, του πιστωτικού κινδύνου και της διαχείρισης του.

Η αλλαγή στις κοινωνικές αντιλήψεις για την πίστωση σηματοδότησαν την γενική αποδοχή της και οδήγησαν κατά συνέπεια στην διάδοση και

επέκτασή της ως έναν τρόπο ζωής και δράσης τόσο στην ιδιωτική ανθρώπινη ζωή όσο και στην επιχειρησιακή πραγματικότητα.

Η αλλαγή στις γεωπολιτικές και οικονομικές συνθήκες οδήγησε σε απελευθεροποίηση και παγκοσμιοποίηση των αγορών, αυξημένη ρευστότητα και οξύμενο ανταγωνισμό σε όλες τις οικονομικές δραστηριότητες των χρηματοοικονομικών ιδρυμάτων. Οι πιέσεις για παροχή πιστώσεων με ελαστικούς όρους και σε μεγάλο αριθμό αιτούμενων αγνώστου ταυτότητος και πιστωτικού παρελθόντος αυξήθηκαν ασφυκτικά. Ο πιστωτικός κίνδυνος όμως έγινε κατ' αυτόν τον τρόπο πολύπλοκος και η διαχείρισή του μία δύσκολη και ιδιαίτερα ευαίσθητη υπόθεση. Μέχρις ώτου όμως αυτό να γίνει αντιληπτό, η επιπόλαιη διαχείρισή του και η έλλειψη μιας σοβαρής και αποτελεσματικής πιστωτικής πολιτικής από την πλευρά των τραπεζών, προκάλεσε πλήθος αθετήσεων και κατέληξε σε οξείες τραπεζικές κρίσεις.

Από την αλλαγή λοιπόν και εξέλιξη της πίστωσης και του πιστωτικού κινδύνου, φτάνουμε αναπόφευκτα στην αναγκαιότητα για την αλλαγή και στη διαχείρισή του. Ειδικά τα παραδείγματα αθετήσεων που παραθέσαμε επισημαίνουν τρανταχτά την ανάγκη για μια μεταβολή στον τρόπο αντίληψης και αντιμετώπισης αυτού του είδους κινδύνου.

1.5.ΕΞΕΛΙΞΕΙΣ ΣΤΗ ΔΙΑΧΕΙΡΙΣΗ ΤΟΥ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

Οι τραπεζικές κρίσεις επομένως των τελευταίων είκοσι χρόνων, αντανakλούν εκτός των άλλων μία κατάρρευση των παραδοσιακών τεχνικών πιστωτικής ανάλυσης. Οι τράπεζες δάνειζαν σε αγορές με πολύ μεγαλύτερο βαθμό κινδύνου από αυτόν που προέβλεπαν οι παραπάνω μεθοδολογίες. Επιπλέον, η παραδοσιακή προσέγγιση του πιστωτικού κινδύνου, διακρινόταν από μία τάση να προκαλεί συγκέντρωση του κινδύνου ή έστω να αγνοεί προκλητικά τις σημαντικές συσχετίσεις μεταξύ διαφορετικών δανείων.

Στον σύγχρονο κόσμο, η πιστωτική ανάλυση καθεμιας περίπτωσης ξεχωριστά δεν είναι πλέον αρκετή. Ο μόνος τρόπος για την εξασφάλιση των τραπεζών έναντι αυτού του είδους κινδύνου, είναι να υιοθετήσουν μία πιο επιτηδευμένη προσέγγιση του προβλήματος, επικεντρώνοντας πλέον σε ομάδες πιστώσεων όπως είναι τα χαρτοφυλάκια, τυποποιώντας τους όρους και την τιμολόγηση των δανείων και συλλέγοντας ιστορικά στοιχεία πιστωτικών συμπεριφορών και αθετήσεων. Η ανάγκη επομένως για μία έγκαιρη αναπροσαρμογή των τεχνικών πιστωτικής ανάλυσης διαφαίνεται επιτακτική.

Παρόλο που συγκριτικά τουλάχιστον με άλλα χρηματοοικονομικά μεγέθη και διαφορετικούς κινδύνους (π.χ. market risk), λίγη αναλυτική εργασία έχει γίνει πάνω στον πιστωτικό κίνδυνο,τελευταία, όλο και καινούργια αναλυτικά εργαλεία για την μέτρηση, τη διαχείριση και τον

έλεγχο του γεννιούνται και προωθούνται. Οι σημερινές τράπεζες βασίζονται πλέον ελάχιστα θα λέγαμε στις πελατειακές τους σχέσεις και πολύ περισσότερο στην πιστωτική ανάλυση που διεξάγουν για κάθε πελάτη - υποψήφιο δανειζόμενο με τη σημαντική αρωγή της τεχνολογίας των πληροφοριών και των καινούργιων μοντέλων και αναλυτικών εργαλείων που έχουν στα χέρια τους.

Πολλές και ποιοτικές πληροφορίες για τον πιστωτικό κίνδυνο ακόμη δεν είναι διαθέσιμες και σοβαρή ακαδημαϊκή έρευνα που να αφορά όλες τις πτυχές του μόλις έχει αρχίσει. Η σύγχρονη κοινωνία παρολαυτά, είναι επαρκώς εξοπλισμένη για να ανταποκριθεί σε μια τέτοια πρόκληση. Η τεχνολογία της πληροφορικής και τα σχετικά με αυτήν αναλυτικά εργαλεία έχουν αναπτυχθεί με αξιοσημείωτο ρυθμό τα τελευταία είκοσι χρόνια. Έχουμε πλέον τη δυνατότητα να συλλέξουμε, να αναλύσουμε, να συγκρίνουμε και να ερμηνεύσουμε πληροφορίες πολύ πιο γρήγορα από τις προηγούμενες γενιές. Η διαθεσιμότητα αυτής ακριβώς της τεχνολογίας επέτρεψε στους δανειστές τα τελευταία χρόνια να παρέχουν πίστωση τόσο εκτέταμενα.

Με την αρωγή των παραπάνω μέσων, ο πιστωτικός κίνδυνος αντιμετωπίζεται πλέον σήμερα όχι ως ένα ανησυχητικό πρόβλημα που απαιτεί τη λήψη αμυντικών μετρων, αλλά ως μία επιθετική ευκαιρία κερδοφορίας. Δεν αποτελεί πια ένα κόστος που συνεπάγεται η τραπεζική δουλειά αλλά κάτι με αξία που μπορεί να γίνει μέχρι και αντικείμενο εμπορίου. Αξιολογικοί οίκοι, εγγυητές και μια ποικιλία από εταιρίες ειδικών αποστολών χρησιμεύουν σήμερα ως σημαντικοί κρίκοι στην αλυσίδα του πιστωτικού κινδύνου.

Μερικές νέες τακτικές διαχείρισης πιστωτικού κινδύνου έχουν ήδη αρχίσει να εφαρμόζονται και μερικές άλλες βρίσκονται ακόμη υπό εκτενή και προσεκτική μελέτη. Ιδρύματα όπως η Chase, Bank of America, Citibank έχουν αναπτύξει παγκόσμια πληροφοριακά συστήματα ελέγχου πιστωτικού κινδύνου που ενημερώνονται συνεχώς έτσι ώστε η έκθεση στον πιστωτικό κίνδυνο και η τιμολόγησή του να ελέγχονται σε πραγματικό χρόνο. Πολλά πιστωτικά ιδρύματα επίσης έχουν αγγίξει ένα στάδιο εξέλιξης, όπου δεν επιθυμούν ούτε χρειάζονται πλέον να κρατούν τα δάνεια μέχρι το τέλος της ζωής τους. Οι λόγοι περιλαμβάνουν αναμφισβήτητα πιέσεις από τους ρυθμιστές αυτών των δανείων καθώς και την εμφάνιση δυναμικών δανειακών αγορών. Επιπλέον, οι τράπεζες έχουν την τάση να προσανατολίζονται προς μια μετατόπιση της έκθεσής τους στον πιστωτικό κίνδυνο μέσω συναλλαγών με άλλους συμβαλλομένους. Οι προαναφερθέντες τραπεζικές κρίσεις τέλος της Λατινικής Αμερικής και των εταιριών real estate, οδήγησαν σε συμπεράσματα για την σημαντικότητα της συσχέτισης του πιστωτικού κινδύνου και την αναγκαιότητα της διαφοροποίησης του μέσω της διαχείρισης δανειακών χαρτοφυλακίων.

Όσο για την αγορά τραπεζικών δανείων, αυξάνεται διαρκώς σε μέγεθος και ρευστότητα.

Απομακρυνόμενες λοιπόν από την παραδοσιακή άποψη ότι η διαχείριση του πιστωτικού κινδύνου είναι κυρίως τέχνη, πολλές τράπεζες υιοθετούν συνεχώς νέες τακτικές. Προσπαθούν να συλλέξουν τόσο τις πληροφορίες όσο και το αναλυτικό υπόβαθρο που είναι απαραίτητα για την αποτίμηση των δανείων σε όρους κινδύνου-απόδοσης. Παρολαυτά, η ανάπτυξη των νέων αυτών τεχνικών υπήρξε ανομοιογενής με αρκετές ελλείψεις και κενά να παραμένουν.

Τα εργαλεία για τη διαχείριση του πιστωτικού κινδύνου δεν επαρκούν από μόνα τους για την εξασφάλιση των χρηστών τους. Κάθε μοντέλο και εργαλείο είναι ένα γέννημα της ανθρώπινης διάνοησης που προσπαθεί να μοντελοποιήσει τον πραγματικό κόσμο μέσω μιας περιορισμένης ομάδας μεταβλητών. Αυτά τα εργαλεία χρησιμοποιούν εκτιμητικές μεθόδους και μεθόδους βελτιστοποίησης που αποτελούν επίσης επινοήσεις του ανθρώπινου μυαλού. Κατ'αυτόν τον τρόπο, ένα μοντέλο μπορεί να αιχμαλωτίσει ένα μεγάλο μέρος της πραγματικότητας αλλά παρολαυτά να παραλείψει και κάποιες σημαντικές πτυχές της. Τα νέα χρηματοοικονομικά εργαλεία είναι με άλλα λόγια έργα υπό εξέλιξη - χρήσιμα αλλά όχι τέλεια. Αν χρησιμοποιηθούν δίχως την ανάλογη προσοχή και κρίση, μπορούν να οξύνουν, όχι να αμβλύνουν την έκθεση ενός πιστωτικού ιδρύματος στον πιστωτικό κίνδυνο. Η αποτελεσματικότητα των παραπάνω εργαλείων εξαρτάται από τις ικανότητες, τα κίνητρα και τη νοοτροπία των ανθρώπων που τα χρησιμοποιούν. Οι συμμετέχοντες με άλλα λόγια στις χρηματαγορές και οι χρήστες αυτών των εργαλείων οφείλουν να είναι ιδιαίτερα προσεκτικοί όσον αφορά τη μοντελοποίηση του κινδύνου.

2. ΠΙΣΤΩΤΙΚΗ ΑΝΑΛΥΣΗ

2.1. ΔΙΑΚΡΙΣΗ ΜΕΤΑΞΥ ΠΟΙΟΤΙΚΗΣ ΚΑΙ ΠΟΣΟΤΙΚΗΣ

Η παραδοσιακή πιστωτική ανάλυση επικεντρώνεται σε δύο ξεχωριστά αλλά αλληλένδετα θέματα: την επιθυμία του δανειζομένου να αποπληρώσει το δάνειο και την ικανότητά του.

Η ανάλυση της επιθυμίας για αποπληρωμή αναφέρεται κυρίως στην έρευνα των ιδιαίτερων χαρακτηριστικών του δανειζομένου και είναι γνωστή ως ποιοτική πιστωτική ανάλυση. Η εκτίμηση από την άλλη της ικανότητας ενός ατόμου ή μιας επιχείρησης να ξεπληρώσει ένα δάνειο, πρόκειται για μια διαφορετική διαδικασία η οποία κοινώς ονομάζεται ποσοτική πιστωτική ανάλυση. Στη δεύτερη αυτή περίπτωση, οι τραπεζίτες συνήθως ξεκινούν την ανάλυση από μακροοικονομικά δεδομένα εξειδικεύοντας διαδοχικά σε μικροοικονομικά στοιχεία που αφορούν τη συγκεκριμένη επιχείρηση. Ερευνούν για παράδειγμα τις προοπτικές της οικονομίας και τον οικονομικό κύκλο στον οποίο βρίσκεται, τις αντίστοιχες προοπτικές του κλάδου της επιχείρησης και την οικονομική της φάση, την οικονομική κατάσταση της συγκεκριμένης επιχείρησης και τις προοπτικές ανάκαμψης ή κερδοφορίας της.

Και τα δύο είδη ανάλυσης απαιτούν πληροφορίες, οι οποίες όσο περισσότερες, ακριβέστερες και έγκαιρες είναι, τόσο σωστότερη θα είναι και η απόφαση για χορηγία δανείου που θα ληφθεί. Παρόλο που και οι δύο τύποι έρευνας είναι κρίσιμοι για την πιστωτική ανάλυση, τα οικονομικά στοιχεία είναι τα πιο σημαντικά. Ιστορικά, περισσότερα χρήματα έχουν χαθεί λόγω της ανικανότητας των δανειζομένων να τα επιστρέψουν παρά λόγω της απροθυμίας τους. Γι' αυτόν ακριβώς το λόγο, στη συγκεκριμένη εργασία θα ασχοληθούμε με την πιστωτική ανάλυση που αφορά τα ποσοτικά οικονομικά στοιχεία και θα αναφερθούμε διεξοδικά στις πρόσφατες εξελίξεις που τη χαρακτηρίζουν.

2.2. ΠΟΣΟΤΙΚΗ ΠΙΣΤΩΤΙΚΗ ΑΝΑΛΥΣΗ

- ΜΕΤΑΣΤΡΟΦΗ ΤΗΣ ΑΝΑΛΥΣΗΣ ΑΠΟ ΤΟΥΣ ΙΣΟΛΟΓΙΣΜΟΥΣ ΣΤΙΣ ΚΑΤΑΣΤΑΣΕΙΣ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ

Η κλασική πιστωτική ανάλυση αναπτύχθηκε σε σχέση με τις θεμελιώδεις αλλαγές που έλαβαν χώρα στις τραπεζικές εργασίες. Ιστορικά, η πρωταρχική αποστολή των τραπεζών ήταν η παροχή κεφαλαίων και η χρηματοδότηση του εμπορίου. Τα δάνεια που παρέχονταν σε αυτές τις περιπτώσεις, ασφαλιζόνταν με πάγια περιουσιακά στοιχεία ή άλλες μορφές αποδεκτών εγγυήσεων. Η απόφαση επομένως για την χορήγηση ή όχι ενός δανείου εξαρτιόταν στο μεγαλύτερο βαθμό από την επάρκεια της αξίας των προτεινόμενων κάθε φορά εγγυήσεων. Επιπλέον, οι τράπεζες παραδοσιακά, επικεντρώνονταν σε βραχυχρόνια δάνεια, το πολύ ενός έτους, ίσα-ίσα για την κάλυψη των εποχιακών αναγκών των πελατών τους. Η βραχυχρόνια αυτή διάσταση των τότε δανείων και η έμφαση στις εγγυήσεις, οδηγούσε αναπόφευκτα στην

επικέντρωση της ανάλυσης στους στατικούς ισολογισμούς των εταιριών και στην πλήρη παράβλεψη των καταστάσεων αποτελεσμάτων και ταμειακών ροών.

Τα τελευταία πενήντα χρόνια όμως, οι τράπεζες προχώρησαν στη χρηματοδότηση και πάγιων στοιχείων των πελατών τους. Η μεταστροφή αυτή μετέβαλλε την χρήση εγγυήσεων σε διαδικασία λιγότερο χρήσιμη και σχετική με την πιστωτική ανάλυση. Δεν υπάρχει άλλωστε ρευστή αγορά εγγυήσεων για στοιχεία όπως είναι οι εξοπλισμοί νοσοκομείων ή τα συστατικά μέρη ενός αυτοκινήτου. Επιπλέον, στις αρχές του 1950, οι τράπεζες άρχισαν να καλύπτουν τις ανάγκες των πελατών τους για πιο μακροχρόνια δάνεια.

Από τη στιγμή λοιπόν που το χρέος έπρεπε να ξεπληρωθεί σε μετρητά, οι τραπεζίτες σταδιακά αναγνώρισαν ότι έπρεπε να επικεντρωθούν πάνω απ' όλα στις ταμειακές ροές του δανειζομένου. Οι ταμειακές ροές που προέρχονται από τις λειτουργίες της επιχείρησης καθορίζουν τη ρευστότητά της, την ικανότητά της δηλαδή να παράγει μετρητά από τις εσωτερικές της λειτουργίες επαρκή για να ικανοποιήσει το χρέος της. Αντιπροσωπεύουν επίσης την αποτελεσματικότητα της διοίκησης να ελέγχει την εισροή και χρήση μετρητών. Από την πλευρά του δανειστή, επαληθεύουν την ικανότητα του δανειζομένου να εξυπηρετεί τις σταθερές του υποχρεώσεις, μέσα στις οποίες περιέχονται και οι επιστροφές δανειακών κεφαλαίων καθώς και οι πληρωμές τόκων. Τέλος, οι καταστάσεις των ταμειακών ροών αντανακλούν την μεταβλητότητα των μεγεθών αυτών μέσα στο χρόνο, μας δίνουν δηλαδή τις πιθανοτικές κατανομές των μελλοντικών ταμειακών ροών και βοηθούν στην εκτίμηση της πιστοληπτικής ικανότητας. Αντικείμενο μελέτης πλέον, άρχισαν να αποτελούν οι καταστάσεις των ταμειακών ροών του υποψήφιου δανειζομένου. Αντίθετα, ο ισολογισμός δεν παρέχει καμία πληροφορία για την ικανότητα της επιχείρησης να παράγει κέρδη. Η πλειοψηφία των μετρητών που θα δεχτεί η επιχείρηση μέσα στους επόμενους δώδεκα μήνες καθώς και η πλειοψηφία των υποχρεώσεων που θα προκύψουν στο ίδιο χρονικό διάστημα, δεν αντιπροσωπεύονται από τα περιουσιακά στοιχεία του ισολογισμού την ημέρα της δημοσίευσής του. Επιπλέον, οι πωλήσεις αποτελούν την κύρια πηγή εισροής μετρητών και ο ισολογισμός δεν μας λέει σχεδόν τίποτα γι' αυτές. Ο δανεισμός με βάση τις ταμειακές ροές αντικατέστησε τον δανεισμό που στηριζόταν στις εγγυήσεις και έλαβε την πρώτη θέση στις δραστηριότητες μιας εμπορικής τράπεζας.

Παρατηρείται λοιπόν σταδιακά μία σαφής στροφή πρώτα από την ποιοτική προς την ποσοτική ανάλυση και στη συνέχεια από την ποσοτική πιστωτική ανάλυση βάσει ισολογισμών (που αποτελούν στατικές καταστάσεις) στην ανάλυση βάσει των ταμειακών ροών (που αποτελούν δυναμικές καταστάσεις).

-ΑΝΑΛΥΣΗ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ ΜΕ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΥΣ ΔΕΙΚΤΕΣ

Η ανάλυση αυτή των ταμειακών ροών γίνεται κατεξοχήν με τη βοήθεια χρηματοοικονομικών δεικτών. Κάθε χρηματοοικονομικός δείκτης εμπεριέχει πληροφορίες τόσο σε απόλυτη όσο και σε συγκριτική βάση. Το πλήθος των χρηματοοικονομικών δεικτών διακρίνεται σε διάφορες κατηγορίες:

- αριθμοδείκτες ρευστότητας (liquidity ratios): προσδιορίζουν τη βραχυχρόνια οικονομική κατάσταση της επιχείρησης και την ικανότητά της να ανταποκρίνεται στις βραχυπρόθεσμες υποχρεώσεις της. Ουσιαστικά, μετρούν την ποιότητα και την ικανότητα των περιουσιακών στοιχείων της επιχείρησης να καλύπτουν τις λήγουσες βραχυπρόθεσμες υποχρεώσεις της. Η ρευστότητα της εταιρίας αποτελεί το πρώτο οικονομικό στοιχείο που λαμβάνουν υπόψη τους οι δανειστές. Τυπικά, μελετούν κατά πόσον τα βραχυπρόθεσμα εποχιακά δάνεια θα μπορέσουν να καλυφθούν από τις βραχυχρόνιες πηγές κεφαλαίων. Τα αποδεκτά όρια ρευστότητας εξαρτώνται από τις εκάστοτε συνθήκες που επικρατούν στην οικονομία και στον συγκεκριμένο κλάδο καθώς και την προβλεψιμότητα των ταμειακών ροών. Εταιρίες που εμφανίζονται διαρκώς συνεπείς στις βραχυχρόνιες υποχρεώσεις τους, αυξάνουν τις επιλογές τους για εύρεση πρόσθετων πηγών χρηματοδότησης. Ενδεικτικά, αναφέρουμε κάποιους αριθμοδείκτες ρευστότητας:
 $(\text{Διαθέσιμο} + \text{απαιτήσεις} + \text{Αποθέματα}) / \text{Βραχυχρόνιες Υποχρεώσεις}$,
 $\text{Διαθέσιμο Ενεργητικό} / \text{Ληξιπρόθεσμες Υποχρεώσεις}$.
- αριθμοδείκτες δραστηριότητας (activity ratios): μετρούν τον βαθμό αποτελεσματικότητας της επιχείρησης στη χρησιμοποίηση των περιουσιακών της στοιχείων. Αποκαλούνται συχνά και asset management ratios και υποθέτουν μια ισορροπία μεταξύ του ενεργητικού και των πωλήσεων της επιχείρησης, η οποία προσδιορίζει πόσο αποτελεσματικά χρησιμοποιεί η εταιρία τις πηγές κεφαλαίων της. Ενδεικτικά αναφέρουμε μερικούς από τους αριθμοδείκτες αυτής της κατηγορίας που χρησιμοποιούνται συχνότερα: Καθαρές πωλήσεις/Μ.Ο. Απαιτήσεων, Αγορές/Μέσο Ύψος Βραχυπρόθεσμων Υποχρεώσεων, Καθαρές Πωλήσεις/Καθαρό Κεφάλαιο Κινήσεως.
- αριθμοδείκτες αποδοτικότητας (profitability ratios): δείχνουν την αποδοτικότητα της επιχείρησης, τη δυναμικότητα των κερδών της και την ικανότητα και επιτυχία της διοίκησης. Παρόλο που ο ύψιστος στόχος μιας εταιρίας είναι η μεγιστοποίηση του πλούτου των μετόχων, η αποδοτικότητα ή αλλιώς κερδοφορία, αποτελεί ένα παραδοσιακό standard για την επιτυχία. Οι δείκτες αποδοτικότητας αποκαλύπτουν σε μεγάλο βαθμό στοιχεία που αφορούν τις πολιτικές και τις αποφάσεις της εταιρίας δείχνοντας έτσι την ικανότητα της διοίκησης να ελέγχει τα έξοδα, να βελτιώνει την αποδοτικότητα και να αντιμετωπίζει προβλήματα

και υφέσεις καθώς και την αποτελεσματικότητά της στον τομέα των επενδύσεων και πωλήσεων της εταιρίας. Οι σημαντικότεροι από αυτούς είναι: Μικτά Κέρδη/Καθαρές πωλήσεις, Κόστος Πωληθέντων/Καθαρές Πωλήσεις, Καθαρά Κέρδη/Σύνολο Ενεργητικού.

- αριθμοδείκτες μόχλευσης (leverage ratios): με αυτούς εκτιμάται η μακροχρόνια ικανότητα της επιχείρησης να ανταποκρίνεται στις υποχρεώσεις της και ο βαθμός προστασίας που αναλαμβάνουν οι πιστωτές της. Η μόχλευση αποτελεί έναν σημαντικό παράγοντα για τον προσδιορισμό της υγείας της επιχείρησης. Σε αντίθεση με τους δείκτες ρευστότητας και αποδοτικότητας, η μόχλευση δεν αποτελεί ένα κατ'ανάγκη μέγεθος προς μεγιστοποίηση, ακόμη κι όταν αυτή μπορεί να αυξήσει την αποδοτικότητα του ενεργητικού ή να ωφελήσει με μεγαλύτερες φορολογικές απαλλαγές. Αντιθέτως, η πρόκληση της χρηματοοικονομικής μόχλευσης συνίσταται στην εξεύρεση της ισορροπίας μεταξύ της ωφέλειας του χρέους και του κόστους του κινδύνου που αυτό περικλείει. Τη διοίκηση απασχολεί το μέγεθος της εξωτερικής χρηματοδότησης. Όσο για τους δανειστές, ενδιαφέρονται για το βαθμό της μόχλευσης που υιοθετούν οι δανειζόμενοι, γιατί δείχνει την έκθεση κινδύνου της επιχείρησης στις πληρωμές των τοκομεριδίων. Όσο περισσότερα τα ξένα κεφάλαια της επιχείρησης τόσο μεγαλύτερη η πιθανότητα να μην μπορέσει να ικανοποιήσει τις απαιτήσεις όλων των πιστωτών της. Επιπλέον, μία εταιρία χρηματοδοτούμενη σε μεγάλο βαθμό από ξένα κεφάλαια, παρέχει στους πιστωτές της μικρότερη προστασία σε περίπτωση χρεωκοπίας. Οι κυριότεροι δείκτες μόχλευσης είναι: Ίδια Κεφάλαια/Συνολικά Κεφάλαια, Ίδια Κεφάλαια/Ξένα Κεφάλαια, Ίδια Κεφάλαια/Πάγια, Καθαρά Κέρδη προ φόρων και τόκων/Σύνολο Τόκων.
- αριθμοδείκτες επενδύσεων (investment or valuation ratios): συσχετίζουν τον αριθμό μετοχών μιας επιχείρησης και τη χρηματιστηριακή τους τιμή με τα κέρδη, τα μερίσματα και τα άλλα περιουσιακά της στοιχεία. Από τη στιγμή που ο Ισολογισμός και οι οικονομικές καταστάσεις διάθεσης των αποτελεσμάτων δεν παρέχουν μια πλήρη εικόνα για την οικονομική κατάσταση της επιχείρησης, απαιτούνται εναλλακτικά μέσα. Μέσα από την ανάλυση των valuation ratios λαμβάνουμε μια ένδειξη της αξίας και του κινδύνου της επιχείρησης. Αν οι ενδείξεις της αγοράς είναι υψηλότερες από τις λογιστικές ενδείξεις, η αγορά έχει αμπιστοσύνη στην επιχείρηση. Οι σημαντικότεροι δείκτες επενδύσεων είναι: Σύνολο Ιδίων Κεφαλαίων /Αριθμός μετοχών, Τρέχουσα Χρηματιστηριακή τιμή/Κέρδη ανά μετοχή, Καθαρά Κέρδη χρήσεως/Σύνολο καταβαλλόμενων Μερισμάτων .

Οι πρώτες επίσημες και συγκεντρωτικές μελέτες για την χρεωκοπία των επιχειρήσεων διεξήχθησαν γύρω στο 1930. Μία μελέτη εκείνης της εποχής, “Changes in the financial structure of unsuccessful corporations” των R.F. Smith και A.H.Winakor και αρκετές άλλες αργότερα, κατέληξαν στο συμπέρασμα ότι οι αποτυχημένες εταιρίες επιδείκνυαν σημαντικά διαφορετικούς χρηματοοικονομικούς δείκτες από τις υγιείς. Μία μεταγενέστερη μελέτη του W.H.Beaver : “Financial ratios as predictors of failure” του 1966, ασχολήθηκε με τη σύγκριση μιας λίστας δεικτών χρεωκοπημένων εταιριών και αντιστοιχισμένων υγιών. Τα αποτελέσματα αυτής της μελέτης για πέντε χρόνια προ της χρεωκοπίας, οδήγησαν στο συμπέρασμα ότι η ανάλυση δεικτών μπορεί να χρησιμοποιηθεί στην πρόβλεψη της χρεωκοπίας. Η μελέτη του W.H.Beaver μπορεί να θεωρηθεί ως πρωτοπόρος και απαρχή μιας σειράς ανάλογων μελετών για το ίδιο θέμα.

Οι δείκτες που αποδείχθηκαν σημαντικότεροι για την πρόβλεψη της χρεωκοπίας είναι οι δείκτες ρευστότητας, αποδοτικότητας και μόχλευσης χωρίς όμως διαυγή σειρά προτεραιότητας.

-ΜΕΙΟΝΕΚΤΗΜΑΤΑ ΑΝΑΛΥΣΗΣ ΤΑΜΕΙΑΚΩΝ ΡΟΩΝ ΜΕ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΟΥΣ ΔΕΙΚΤΕΣ

Η κλασική πιστωτική ανάλυση σύντομα αποδείχτηκε ότι παρουσιάζει μία σειρά από σημαντικά προβλήματα.

Καταρχήν, ένα σύστημα εξειδικευμένου προσωπικού όπως αυτό της παραδοσιακής πιστωτικής ανάλυσης, απαιτεί πολύ υψηλό κόστος για να διατηρηθεί. Οι τράπεζες πρέπει να έχουν διαρκώς στη διάθεσή τους αρκετούς ειδικούς για να αντιμετωπίζουν τον διαρκώς αυξανόμενο όγκο της πιστωτικής ανάλυσης καθώς και έναν μεγάλο αριθμό ατόμων που θα εκπαιδεύουν τους παραπάνω ειδικούς. Η υποκειμενική κρίση δηλαδή στην οποία στηρίζεται το σύστημα αυτό, απαιτεί την εκπαίδευση και εργασία πολλών ειδικών, ώστε να αποφευχθούν όσο το δυνατόν περισσότερο τα λάθη.

Η παραπάνω εκπαίδευση βέβαια δεν απομακρύνει τελείως το πρόβλημα του ανθρώπινου παράγοντα. Πολλές φορές, πιστωτικά προβλήματα έχουν προκύψει από ανεπαρκή εκτέλεση της πιστωτικής ανάλυσης. Ένα σύστημα ειδικών είναι τόσο καλό όσο οι ειδικοί που το αποτελούν. Όπως κάθε ανθρώπινο σύστημα, μπορεί να λειτουργήσει με πολύ ακανόνιστους τρόπους. Ένας και μόνο λάθος άνθρωπος μπορεί να θέσει σε κίνδυνο την βιωσιμότητα ολόκληρου του πιστωτικού ιδρύματος.

Ένα ακόμη σημαντικό μειονέκτημα της παραδοσιακής πιστωτικής ανάλυσης, είναι η αδυναμία της για έναν συνεχή και συστηματικό έλεγχο του μεγέθους της πιστωτικής έκθεσης και του κινδύνου του ιδρύματος. Οι πιστωτικές αναλύσεις των ειδικών είναι πολύ συχνά ετεροχρονισμένες τη στιγμή που φτάνουν στις τράπεζες. Οι εξελίξεις σήμερα είναι τόσο ραγδαίες που καμία στατική ανάλυση δεν θα μπορούσε να τις συλλάβει.

Επιπλέον, η κλασική πιστωτική ανάλυση, έχει συχνά οδηγήσει τις τράπεζες σε μία λανθασμένη έννοια ασφαλείας, αποτυγχάνοντας να τις προστατεύσει έναντι των συστηματικών κινδύνων.

Μελετώντας την παραδοσιακή πιστωτική ανάλυση πιο αναλυτικά και επικεντρώνοντας στη χρήση των χρηματοοικονομικών δεικτών, θα συναντήσουμε ένα ακόμη μειονέκτημα που δυσκολεύει κατά πολύ τη χρήση της. Οι μονομεταβλητές αυτές μέθοδοι, προϋποθέτουν την σύγκριση κάποιων χρηματοοικονομικών δεικτών-κλειδιών του υποψήφιου δανειζομένου με κάποια standards που έχουν τεθεί βασιζόμενα στους αντίστοιχους δείκτες-κλειδιά του κλάδου. Κατ' αυτόν τον τρόπο, ο αναλυτής είναι σε θέση να μελετήσει αν και σε ποιο βαθμό ένας συγκεκριμένος δείκτης της επιχείρησης-πελάτη διαφέρει από τον μέσο αντίστοιχο του κλάδου. Στην πραγματικότητα όμως, το μη ικανοποιητικό επίπεδο ενός δείκτη μετριάζεται συχνά και αντισταθμίζεται από τη δύναμη ενός άλλου μέτρου. Μία επιχείρηση για παράδειγμα, μπορεί να έχει φτωχό δείκτη αποδοτικότητας αλλά υψηλό δείκτη ρευστότητας. Ο περιορισμός λοιπόν που μας θέτει η μονομεταβλητή αυτή μέθοδος, είναι η δυσκολία εξαγωγής συμπερασμάτων σε περίπτωση τέτοιων αντιφατικών δεδομένων.

Η υποκειμενικότητα και στατικότητα λοιπόν αυτής της μεθόδου, το υψηλό της κόστος και η αδυναμία της να λαβει υπόψη τις συσχετίσεις των διαφόρων πιστωτικών κινδύνων και των ποικίλων χρηματοοικονομικών δεικτών, οδήγησαν στην επιτακτική ανάγκη για την εύρεση καινούργιων μεθοδολογιών και εργαλείων.

2.3. ΣΤΡΟΦΗ ΑΠΟ ΧΡΗΣΗ ΧΡΗΜΑΤΟΟΙΚΟΝΟΜΙΚΩΝ ΔΕΙΚΤΩΝ ΣΕ ΧΡΗΣΗ ΟΙΚΟΝΟΜΕΤΡΙΚΩΝ ΜΟΝΤΕΛΩΝ (FROM UNIVARIATE TO MULTIVARIATE)

Παρόλο που τα μονομεταβλητά μοντέλα βρίσκονται ακόμη σε χρήση σε αρκετά πιστωτικά ιδρύματα, οι περισσότεροι ακαδημαϊκοί και ένας αυξανόμενος αριθμός στελεχιακού δυναμικού, φαίνεται να αποδοκιμάζουν πλέον την ανάλυση χρηματοοικονομικών δεικτών ως μέσο για τον προσδιορισμό της οικονομικής κατάστασης μιας επιχείρησης. Πολλοί αξιολογοί θεωρητικοί απορρίπτουν τις παραδοσιακές μεθόδους και κλίνουν προς την εφαρμογή αυστηρότερων στατιστικών τεχνικών. Οι τελευταίες αυτές τεχνικές όμως θα έπρεπε να θεωρηθούν ως μία βελτίωση της παραδοσιακής πιστωτικής ανάλυσης δεικτών παρά ως μία ριζική αποχώρηση από αυτήν. Μία κατάλληλη επέκταση λοιπόν των προηγούμενων μελετών έγκειται στη χρήση των ευρημάτων τους και στο συνδυασμό αυτών και κάποιων άλλων μέτρων μέσα σε ένα προβλεπτικό μοντέλο.

3.ΜΟΝΤΕΛΑ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

3.1.ΧΡΗΣΙΜΟΤΗΤΑ ΜΟΝΤΕΛΩΝ

Τα χρηματοοικονομικά μοντέλα αντιπροσωπεύουν μία συσσώρευση των ανθρωπίνων γνώσεων, εμπειριών και πειραμάτων που μπορούν να χρησιμοποιηθούν για την ερμηνεία του τρόπου συμπεριφοράς των ανθρώπων ή λειτουργίας των πραγμάτων. Κάθε μοντέλο και εργαλείο είναι ένα γέννημα της ανθρώπινης διανόησης που προσπαθεί να μοντελοποιήσει τον πραγματικό κόσμο μέσω μιας περιορισμένης ομάδας μεταβλητών και κάτω από ορισμένους περιορισμούς. Ένα μοντέλο διευκολύνει σε μεγάλο βαθμό την κατανόηση ενός φαινομένου και κατά συνέπεια και την αξιοποίησή του. Τα μοντέλα που μετρούν τον πιστωτικό κίνδυνο δεν αποτελούν εξαίρεση. Με ένα μοντέλο πιστωτικού κινδύνου αναζητούμε έμμεσα ή άμεσα απαντήσεις σε ερωτήσεις του είδους: δεδομένης της προηγούμενης εμπειρίας μας και των υποθέσεών μας για το μέλλον, ποια είναι η αξία ενός συγκεκριμένου δανείου και ποιός ο κίνδυνος οι συμφωνημένες ταμειακές ροές να μην πραγματοποιηθούν.

3.2.ΚΑΤΑΣΚΕΥΗ ΜΟΝΤΕΛΩΝ

Η κατασκευή ενός μοντέλου πιστωτικού κινδύνου προϋποθέτει την αρωγή και χρήση πολλών ξεχωριστών συστατικών στοιχείων. Πρώτα καθορίζουμε τις μεταβλητές που προσδιορίζουν και επηρεάζουν τον πιστωτικό κίνδυνο. Στη συνέχεια θέτουμε τις σχέσεις που συνδέουν αυτές τις μεταβλητές, χρησιμοποιώντας το θεωρητικό υπόβαθρο που έχουμε στη διάθεσή μας. Κατόπιν, μία ομάδα δεδομένων είναι απαραίτητη, καθώς κανένα μοντέλο δεν μπορεί να δημιουργηθεί στο κενό. Υπάρχουν μάλιστα περιπτώσεις, όπου ο μόνος τρόπος για να ανακαλύψουμε νέες σχέσεις είναι να εργαστούμε με βάση μόνο τα στοιχεία χωρίς να υπάρχει κάποιο θεωρητικό υπόβαθρο. Τέλος, μια σειρά από τεστ πρέπει να διεξαχθούν ώστε να βεβαιωθούμε ότι το μοντέλο λειτουργεί όπως αναμενόταν.

3.3.ΤΕΧΝΙΚΕΣ ΠΟΥ ΧΡΗΣΙΜΟΠΟΙΟΥΝΤΑΙ

Οι πιο συνήθεις τεχνικές που χρησιμοποιούνται είναι οι εξής:

- Οικονομετρικές τεχνικές: είναι η linear και multiple discriminant analysis, η πολλαπλή παλινδρόμηση(multiple regression), η logit analysis και η probit analysis. Όλες οι παραπάνω τεχνικές μοντελοποιούν την πιθανότητα μη αποπληρωμής του δανείου σαν μία εξαρτημένη μεταβλητή, της οποίας η διακύμανση ερμηνεύεται από μία ομάδα ανεξάρτητων μεταβλητών. Οι ανεξάρτητες μεταβλητές περιέχουν χρηματοοικονομικούς δείκτες

καθώς και κάποιες εξωτερικές μεταβλητές που χρησιμοποιούνται για τη μέτρηση των οικονομικών συνθηκών.

- Neural Networks: πρόκειται για συστήματα που βασίζονται στην πληροφορική και που αποσκοπούν στη μίμηση της λειτουργίας του ανθρώπινου εγκεφάλου και συγκεκριμένα των νευρωνίων του που αποτελούν τις μικρότερες μονάδες λήψης αποφάσεων. Χρησιμοποιούν τα ίδια δεδομένα με τις οικονομετρικές μεθόδους, αλλά καταλήγουν στο μοντέλο λήψης αποφάσεων με εναλλακτικές εφαρμογές της μεθόδου δοκιμής σωστού και λάθους (trial and error method).
- Μοντέλα Βελτιστοποίησης (Optimization models): είναι μαθηματικές τεχνικές προγραμματισμού οι οποίες ανακαλύπτουν τους μέγιστους συντελεστές στάθμισης για τον δανειζόμενο και το δάνειο που ελαχιστοποιούν το λάθος του δανειστή και μεγιστοποιούν τα κέρδη.
- Rule-based or expert systems: προσπαθούν να μιμηθούν με έναν δομημένο τρόπο τη διαδικασία που ακολουθεί ένας έμπειρος αναλυτής μέχρι να φτάσει στην πιστωτική απόφαση.
- Hybrid systems using direct computation, estimation and simulation: δημιουργούνται από μία άμεση αιτιώδη σχέση, οι παράμετροι της οποίας καθορίζονται με estimation techniques. Ένα παράδειγμα είναι το μοντέλο KMV που χρησιμοποιεί την θεωρία των options για να ερμηνεύσει την αποτυχία αποπληρωμής.

3.4.ΕΦΑΡΜΟΓΕΣ ΚΑΙ ΠΛΕΟΝΕΚΤΗΜΑΤΑ

Τα χρηματοοικονομικά μοντέλα χρησιμοποιούνται σε μια ποικιλία εφαρμογών όπως είναι η έγκριση και τιμολόγηση των δανείων, η ταξινόμηση και κατάταξη εταιριών ή χρεογράφων ανάλογα με τον πιστωτικό τους κίνδυνο (credit rating), η έγκαιρη προειδοποίηση σε περίπτωση προβλήματος, η επιλογή της σωστής στρατηγικής και αντιμετώπισης του εκάστοτε χρεώστη.

Τα μοντέλα πιστωτικού κινδύνου είναι σημαντικά σήμερα γιατί παρέχουν στον λήπτη των αποφάσεων διορατικότητα και γνώση που δεν θα ήταν εμφανείς σε άλλες περιπτώσεις παρά μόνο ίσως με κάποιο απαγορευτικό κόστος. Μεσα σε μία αγορά, που τα περιθώρια στενεύουν και η πίεση για χαμηλότερα κόστη γίνεται αδυσώπητη, τα μοντέλα παρέχουν στους χρήστες τους ανταγωνιστικό πλεονέκτημα. Σε οποιοδήποτε μεγάλο χρηματοοικονομικό ίδρυμα που διακρίνεται από μία μεγάλη ποικιλία εκθέσεων σε κίνδυνο τα ποσοτικά μοντέλα δίνουν έναν μεγάλο και χρήσιμο βαθμό αντικειμενικότητας.

3.5.ΟΡΙΣΜΟΙ ΜΟΝΤΕΛΩΝ

Στις μονομεταβλητές μεθόδους ανάλυσης πιστωτικού κινδύνου, ο αναλυτής συγκρίνει διάφορους δείκτες-κλειδιά του πιθανού δανειστή με αντίστοιχους πρότυπους του κλάδου. Στα πολυμεταβλητά μοντέλα, οι δείκτες-κλειδιά συνδυάζονται και σταθμίζονται για να παράγουν είτε ένα score πιστωτικού κινδύνου είτε ένα μέτρο της πιθανότητας μη αποπληρωμής. Αν το score πιστωτικού κινδύνου ή η πιθανότητα αγγίζουν μια αξία πάνω από μία κριτική τιμή, ο υποψήφιος δανειζόμενος απορρίπτεται ή τίθεται υπό αυξημένο έλεγχο.

Υπάρχουν τουλάχιστον τέσσερις μεθοδολογικές προσεγγίσεις για την ανάπτυξη πολυμεταβλητών μοντέλων πιστωτικού κινδύνου: α) το linear probability model, β) το logit model, γ) το probit model και δ) το discriminant analysis model. Οι πιο διαδεδομένες σε χρήση είναι το logit και η discriminant analysis.

α)DISCRIMINANT ANALYSIS

Στην linear discriminant analysis προσπαθούμε να βρούμε μια γραμμική συνάρτηση λ'x με k επεξηγηματικές μεταβλητές, που να δίνει τον καλύτερο διαχωρισμό ανάμεσα σε δύο γκρουπ αντιστοιχισμένα σε $y=0$ και $y=1$. Το λ πρέπει να επιλεγθεί έτσι ώστε η διακύμανση του λ'x ανάμεσα στα δύο γκρουπ (variance between groups) να είναι όσο το δυνατόν μεγαλύτερη σε σχέση με τη διακύμανση μέσα σε κάθε γκρουπ (variance within groups). Οι διακυμάνσεις πρέπει να διακρίνονται από το παραπάνω χαρακτηριστικό, ώστε να διευκολύνεται και να γίνεται ακριβέστερη η κατάταξη. Στην περίπτωση για παράδειγμα του πιστωτικού κινδύνου, προσπαθούμε να ελέγξουμε σε ποιο από τα δύο γκρουπ, υγιείς επιχειρήσεις-χρεωκοπημένες επιχειρήσεις, ανήκει μια δεδομένη εταιρία. Για να είναι ακριβής η κατάταξη, θα πρέπει τα χαρακτηριστικά των εταιριών που ανήκουν στο γκρουπ των χρεωκοπημένων να είναι όσο το δυνατόν πιο διαφορετικά από τα χαρακτηριστικά των εταιριών που ανήκουν στο γκρουπ των υγιών. Αυτό ακριβώς σημαίνει η διακύμανση ανάμεσα στα δύο γκρουπ να είναι όσο το δυνατόν μεγαλύτερη. Παράλληλα όμως, θα πρέπει τα χαρακτηριστικά των εταιριών που ανήκουν σε καθένα από τα δύο γκρουπ να είναι όσο το δυνατόν πιο όμοια μεταξύ τους, οι εταιρίες δηλαδή που ανήκουν στο χρεωκοπημένο γκρουπ να διακρίνονται από πανομοιότυπα χαρακτηριστικά και οι υγιείς εταιρίες με τη σειρά τους το ίδιο, ώστε το κάθε γκρουπ να είναι ομοιόμορφο και ομοιογενές. Γι' αυτό προσπαθούμε η διακύμανση μέσα σε κάθε γκρουπ να είναι όσο το δυνατόν μικρότερη. Για να επιτευχθούν λοιπόν τα παραπάνω προσπαθούμε να επιλέξουμε εταιρίες με τέτοια χαρακτηριστικά που να ικανοποιούνται οι δύο παραπάνω προϋποθέσεις για τις διακυμάνσεις.

Το πρόβλημα δηλαδή που προσπαθεί να αντιμετωπίσει η discriminant analysis έχει να κάνει με την κατάταξη ενός ατόμου σε έναν από δύο πληθυσμούς π_1 και π_2 που βασίζονται σε ένα διάνυσμα χαρακτηριστικών $x=(x_1, x_2, x_3 \dots x_k)$.

Η μορφή της discriminant analysis είναι ως εξής:

$$y=\lambda_1x_1+\lambda_2x_2+\lambda_3x_3+\dots+\lambda_kx_k$$

Στην περίπτωση του πιστωτικού κινδύνου δηλαδή, η discriminant analysis προσπαθεί να βρει μια γραμμική συνάρτηση λογιστικών μεταβλητών και μεταβλητών της αγοράς που να ξεχωρίζει όσο το δυνατόν καλύτερα ανάμεσα σε δύο ομάδες κατάταξης δανειζομένων –ομάδα αποπληρωμής και ομάδα μη αποπληρωμής.

Η linear discriminant analysis βασίζεται στις παρακάτω υποθέσεις:

- Οι $f_1(x)$ και $f_2(x)$ είναι πολυμεταβλητές κανονικές.
- Οι πίνακες συνδιακυμάνσεων είναι ίσοι ($\Sigma_1 = \Sigma_2$). Στην περίπτωση που δεν είναι ίσοι ($\Sigma_1 \neq \Sigma_2$) η discriminant function δεν είναι γραμμική αλλά quadratic ως προς το x .
- Οι prior probabilities p_1 και p_2 είναι γνωστές. (Όπου prior probabilities p_1 και p_2 είναι οι αναλογίες των γκρουπ π_1 και π_2 μέσα στον πληθυσμό)
- Οι μέσοι μ_1 και μ_2 των x μέσα στα γκρουπ και ο πίνακας συνδιακυμάνσεων Σ είναι γνωστά.

Πλεονεκτήματα MDA:

- i) Λαμβάνει υπόψη συνολικά τα χαρακτηριστικά που είναι κοινά για τις εταιρίες καθώς και την αλληλοεπίδραση τους.
- ii) Μειώνει τις διαστάσεις του χώρου του αναλυτή, δηλαδή μετασχηματίζει τις ατομικές αξίες των ανεξάρτητων μεταβλητών σε ένα μόνο discriminant score z : εκεί που έπρεπε να αναλύσουμε έναν-έναν τους δείκτες έχουμε τώρα να αναλύσουμε μόνο το z .
- iii) Αναλύει ολόκληρο το profile των μεταβλητών της κάθε εταιρίας ταυτόχρονα και όχι διαδοχικά όπως γινόταν με την univariate approach.

Μειονεκτήματα MDA:

- i) Η discriminant analysis παρουσιάζει μειωμένη ακρίβεια σε μεγάλο δείγμα λόγω των πολλών περιπτώσεων που εισέρχονται. Δεν έχει δηλαδή την ικανότητα να συγκρίνει ανάμεσα σε εταιρίες outstanding, normal και vulnerable παρά μόνο ανάμεσα σε healthy και unsound.

β) LINEAR PROBABILITY MODEL

Ο όρος linear probability model χρησιμοποιείται για να δηλώσει ένα μοντέλο παλινδρόμησης στο οποίο η εξαρτημένη μεταβλητή y είναι μια binary μεταβλητή που λαμβάνει την τιμή 1 αν το γεγονός συμβεί και την τιμή 0 αν δεν συμβεί. Στην περίπτωση δηλαδή του πιστωτικού κινδύνου, η y μπορεί να παίρνει την τιμή 0 αν η εταιρία χρεωκοπήσει και την τιμή 1 αν συνεχίσει να επιβιώνει. Το linear probability model παίρνει την εξής μορφή:

$$y_i = \beta'x_i + u_i$$

με $E(u_i) = 0$ και $E(y_i/x_i) = \beta'x_i$.

Με μια σειρά αλγεβρικών πράξεων καταλήγουμε στην παρακάτω εξίσωση:

$$\text{Var}(u_i) = E(y_i)[1 - E(y_i)]$$

Παρατηρούμε λοιπόν ότι το u_i εξαρτάται από τις τιμές του y_i . Εξαιτίας αυτού του προβλήματος ετεροσκεδαστικότητας ο εκτιμητής ελαχίστων τετραγώνων (Ordinary Least Squares estimator) του β δεν είναι ικανός (efficient).

Επιπλέον, τα κατάλοιπα u_i δεν παρουσιάζουν κανονική κατανομή και γι' αυτό το λόγο η μέθοδος των ελαχίστων τετραγώνων δεν είναι πλήρως αποτελεσματική (fully efficient). Υπάρχουν δηλαδή άλλες μέθοδοι μη γραμμικές περισσότερο αποτελεσματικές. Αυτό το πρόβλημα όμως δεν επηρεάζει τα τεστ σημαντικότητας αν υποθέσουμε ότι οι επεξηγηματικές μεταβλητές έχουν πολυμεταβλητή κανονική κατανομή.

Το σημαντικότερο όμως πρόβλημα του linear probability model είναι ότι η δεσμευμένη πιθανότητα να συμβεί το γεγονός (δεδομένης μιας τιμής του x), μπορεί πρακτικά να βγει εκτός των ορίων (0,1).

Το linear probability model με άλλα λόγια παρουσιάζει κάποια σημαντικά μειονεκτήματα που συνοψίζονται ως εξής:

Μειονεκτήματα:

- i) Ετεροσκεδαστικότητα $\rightarrow \beta_{OLS}$ όχι ικανός εκτιμητής
- ii) Μη κανονική κατανομή των καταλοίπων \rightarrow μέθοδος ελαχίστων τετραγώνων όχι αποτελεσματική
- iii) Το $E(y_i / x_i)$ μπορεί να βγαίνει εκτός των ορίων (0,1)

Πλεονεκτήματα:

- i) Η σκοπιμότητα ύπαρξης όμως του μοντέλου αυτού δικαιολογείται από τη **σχέση** που έχει **με** την **linear discriminant function**. Τα F-tests για να ελέγξουμε αν υπάρχουν σημαντικές διαφορές ανάμεσα στα δύο γκρουπ και t-tests για να ελέγξουμε το $\lambda_i = 0$ στην linear discriminant function και $\beta_i = 0$ στο linear probability model έχουν την ίδια μορφή.
- ii) Με πολλαπλές παρατηρήσεις, μεγάλο δείγμα και με $p_i \neq 1$ ή 0, το linear probability model γίνεται καλύτερο και δεν διαφέρει στο πνεύμα από το logit ή το log-linear. Καλύτερο γίνεται, γιατί το μεγάλο δείγμα μας επιτρέπει να χρησιμοποιήσουμε αντί της ordinary least squares method την weighted least squares method που χρησιμοποιεί ουσιαστικά την minimum chi-square method, η οποία μας δίνει έναν πολύ καλό εκτιμητή για το β , αντιμετωπίζοντας έτσι τα δύο πρώτα προβλήματα του μοντέλου.

γ) PROBIT & LOGIT MODEL

Μία εναλλακτική προσέγγιση που ονομάστηκε από τον Goldberger (1964) probit analysis, είναι να υποθέσουμε ότι υπάρχει μια response variable y_i^* που ορίζεται από τη συνάρτηση παλινδρόμησης :

$$y_i^* = \beta' x_i + u_i$$

Στην πραγματικότητα, η y_i^* δεν είναι παρατηρήσιμη. Αυτό που παρατηρούμε είναι μια dummy variable y που ορίζεται ως εξής:

$$y = 1 \quad \text{αν } y_i^* > 0 \\ y = 0 \quad \text{αλλιώς}$$

Σ' αυτή τη συνάρτηση το $\beta' x_i$ δεν είναι ίσο με $E(y_i / x_i)$ όπως στο linear probability model αλλά ίσο με $E(y_i^* / x_i)$.

Στην προσπάθειά μας να εκτιμήσουμε το β βρίσκουμε την συνάρτηση πιθανοφάνειας (Likelihood function):

$$L = \prod_{y_i=0} F(-\beta' x_i) \prod_{y_i=1} [1 - F(-\beta' x)]$$

Η λειτουργική μορφή για την F θα εξαρτηθεί από τις υποθέσεις που θα κάνουμε για τα u_i . Αν η αθροιστική κατανομή των u_i είναι logistic προκύπτει το **logit** model:

$$F(-\beta' x_i) = 1 / [1 + \exp(\beta' x_i)]$$

Η logit analysis χρησιμοποιεί μια ομάδα λογιστικών μεταβλητών για να προβλέψει την πιθανότητα αποτυχίας του δανειζομένου να ξεπληρώσει το δάνειο, υποθέτοντας ότι αυτή η πιθανότητα κατανέμεται λογαριθμικά. Η αθροιστική πιθανότητα για παράδειγμα μη αποπληρωμής παίρνει μια λογαριθμική μορφή και είναι εξ'ορισμού περιορισμένη να κυμαίνεται μεταξύ των τιμών 0 και 1.

Αν υποθέσουμε ότι τα u_i κατανέμονται κανονικά, $u_i \sim IN(0, \sigma^2)$, τότε παίρνουμε το **probit** model:

$$F(-\beta' x_i) = \int_{-\infty}^{\frac{-\beta' x_i}{\sigma}} \frac{1}{2\sqrt{\pi}} \cdot \exp\left\{-\frac{t^2}{2}\right\} dt$$

Σχέση μεταξύ logit και probit:

Το logit model έχει πιο απλή μορφή από το probit, καθώς δεν περιέχει ολοκληρώματα. Βέβαια, με τη χρήση των σύγχρονων ηλεκτρονικών υπολογιστών, οι υπολογισμοί πολύπλοκων συναρτήσεων με ολοκληρώματα δεν αποτελούν πλέον πρόβλημα ούτε θέτουν περιορισμούς στη χρήση των μοντέλων. Επομένως, το συγκριτικό αυτό πλεονέκτημα του logit έχει πλέον πολύ μικρή σημασία.

Τα δύο αυτά μοντέλα έχουν το πλεονέκτημα ότι δεν παρουσιάζουν αποτελέσματα για το y_i εκτός των ορίων (0,1).

Εκτός της περίπτωσης όπου τα δείγματα είναι πολύ μεγάλα και υπάρχουν πολλές παρατηρήσεις στις tails, στις υπόλοιπες περιπτώσεις, τα δύο μοντέλα παράγουν πολύ όμοια αποτελέσματα. Αυτό συμβαίνει γιατί η αθροιστική κανονική κατανομή του probit και η logistic κατανομή του logit μοιάζουν πολύ μεταξύ τους. Παρολαυτά, οι εκτιμητές του β των δύο μεθόδων δεν είναι άμεσα συγκρίσιμοι. Επειδή η logistic distribution παρουσιάζει διακύμανση $\pi^2/3$, οι εκτιμητές του β που λαμβάνουμε από το logit model πρέπει να πολλαπλασιαστούν με $3^{1/2}/\pi$ για να γίνουν συγκρίσιμοι με τους εκτιμητές του probit. Ο Amemiya πρότεινε οι εκτιμητές του logit να πολλαπλασιαστούν με το $1/1,6 = 0,625$ αντί του $3^{1/2}/\pi$ για μεγαλύτερη προσέγγιση. Παρατήρησε επίσης ότι οι εκτιμητές των logit και linear probability μοντέλων συνδέονται με τις παρακάτω σχέσεις :

Σχέση μεταξύ logit και linear probability:

$$\begin{array}{ll} \beta_{LP} \sim 0,25\beta_L & \text{αν δεν υπάρχει σταθερός όρος} \\ \text{και } \beta_{LP} \sim 0,25\beta_L + 0,5 & \text{αν υπάρχει σταθερός όρος} \end{array}$$

Σχέση μεταξύ probit και linear probability:

$$\beta_P \sim 2,5\beta_{LP} + (c - 1,25) \quad , \text{ όπου } c = \text{σταθερός όρος}$$

Σχέση μεταξύ logit και discriminant analysis:

Αν οι ανεξάρτητες μεταβλητές κατανέμονται κανονικά ο εκτιμητής της discriminant analysis είναι ο εκτιμητής μεγίστης πιθανοφάνειας και επομένως είναι ασυμπτωτικά πιο αποτελεσματικός από τον εκτιμητή μεγίστης πιθανοφάνειας του logit. Αν όμως οι ανεξάρτητες μεταβλητές δεν κατανέμονται κανονικά ο εκτιμητής της discriminant analysis δεν είναι καν συνεπής, ενώ αντίθετα ο εκτιμητής μεγίστης πιθανοφάνειας του logit είναι συνεπής και επομένως πιο δυνατός.

3.6.ΣΥΝΟΠΤΙΚΗ ΑΝΑΦΟΡΑ ΣΤΙΣ ΕΜΠΕΙΡΙΚΕΣ ΜΕΛΕΤΕΣ ΤΟΥ ΠΑΡΕΛΘΟΝΤΟΣ

Τις τελευταίες δεκαετίες αρκετές μελέτες έχουν γίνει με θέμα την χρεωκοπία των εταιριών. Οι πρώτες απόπειρες στις Ηνωμένες Πολιτείες να χρησιμοποιηθούν διαθέσιμα στοιχεία με ποικίλες στατιστικές τεχνικές για την πρόβλεψη της εταιρικής αποτυχίας έγιναν από τον Beaver το 1967 και τον Altman το 1968.

Ο **Beaver** ήταν ο πρώτος που μελέτησε τα χαρακτηριστικά των αποτυχημένων εταιριών σε σύγκριση με αυτά ενός αντιστοιχισμένου δείγματος υγιών. Χρησιμοποιώντας μονομεταβλητές μεθόδους (univariate discriminating tests), ανακάλυψε ότι κάποιοι χρηματοοικονομικοί δείκτες μπορούν να χρησιμοποιηθούν ως χρήσιμοι προβλεπτές της χρεωκοπίας ακόμη και 5 χρόνια πριν αυτή συμβεί. Η μελέτη αυτή μπορεί να θεωρηθεί ως η πρωτοπόρος για το συγκεκριμένο θέμα δίνοντας μάλιστα το έναυσμα για την περαιτέρω έρευνα που ακολούθησε.

Στη συνέχεια της πρώτης αυτής μελέτης, έγινε αισθητό ότι η στήριξη των προβλεπτικών μοντέλων χρεωκοπίας μόνο σε δείκτες ήταν πολύ απλή για να συλλάβει την πολυπλοκότητα της χρηματοοικονομικής αποτυχίας. Αποτέλεσμα αυτής της παραδοχής ήταν η εφαρμογή δύο σημαντικών στατιστικών τεχνικών, της Multiple Discriminant Analysis (MDA) και Regression Analysis (RA), από πολλούς αναλυτές για την πρόβλεψη επικείμενων χρεωκοπιών. Ο **E.Altman** (1968, 1978) ήταν ο πρώτος που χρησιμοποίησε την MDA μέθοδο στο συγκεκριμένο πρόβλημα, δημιουργώντας έτσι ένα μοντέλο γνωστό ως Z-Altman score που αποδείχτηκε κατά 90% ακριβές στην κατάταξη των εταιριών σε υγιείς και μη, έναν χρόνο πριν την αποτυχία. Εννιά χρόνια αργότερα, οι **Altman**, **Haldeman** και **Narayanan** μελέτησαν την προβλεπτική ικανότητα ενός μοντέλου discriminant analysis με 7 μεταβλητές. Το μοντέλο αυτό, γνωστό ως ZETA model βελτιώνει στην ουσία το προηγούμενο μοντέλο του Altman, Altman z-score με τις 5 μεταβλητές.

Ο **Scott** (1981) συγκρίνει πολλά από αυτά τα εμπειρικά μοντέλα με μία θεωρητική προσέγγιση του προβλήματος και συμπεραίνει ότι το ZETA model προσεγγίζει περισσότερο από κάθε άλλο την θεωρητική του δομή για το θέμα της χρεωκοπίας.

Η MDA εφαρμόστηκε επίσης και από τον **Deakin** (1972) ο οποίος βρήκε ότι τα μοντέλα του ήταν τουλάχιστον κατά 95% ακριβή για τα πρώτα τρία έτη πριν τη χρεωκοπία.

Η Regression Analysis εφαρμόστηκε από τον **Edmister** (1971) ο οποίος πέτυχε αποτελέσματα υψηλής ακρίβειας όσον αφορά την κατάταξη.

Οι δύο τεχνικές συγκρίθηκαν σε μία μελέτη του **Collins** (1980), ο οποίος κατέληξε ότι και οι δύο παρείχαν ικανοποιητικά προβλεπτικά αποτελέσματα. Όταν χρησιμοποιούνταν στην ίδια ομάδα δεδομένων η MDA απέδιδε το ίδιο ή και καλύτερα από την RA.

Ο **Martin** (1977) χρησιμοποίησε και την logit και την discriminant analysis προκειμένου να προβλεψει χρεωκοπίες τραπεζών για την περίοδο 1975-1976,

κατά την οποία 23 τράπεζες χρεωκόπησαν. Και τα δύο μοντέλα έδωσαν όμοιες κατατάξεις σε όρους αναγνώρισης χρεωκοπίας/μη χρεωκοπίας.

Ο **West** (1985) χρησιμοποίησε το logit model για να μετρήσει την χρηματοοικονομική κατάσταση των πιστωτικών ιδρυμάτων και να αντιστοιχίσει σε αυτά την πιθανότητα να είναι προβληματικές τράπεζες.

Ο **Lawrence**(1992) εφάρμοσε το logit model για να προβλέψει την πιθανότητα μη αποπληρωμής σε στεγαστικά δάνεια. Ανακάλυψε ότι η ιστορία της πιστωτικής συμπεριφοράς είναι ο πιο σημαντικός προβλεπτικός παράγοντας της αθέτησης.

Σε άλλες χώρες τώρα, η στατιστική μέθοδος που χρησιμοποιήθηκε ήταν η MDA. Στην Ιαπωνία αρκετές μελέτες (π.χ. Nikkei – Business, Takahashi and Ko) κατέληξαν σε υψηλής ακρίβειας κατάταξης αποδόσεις (85% ή περισσότερο). Κάποιες άλλες μελέτες που εφάρμοσαν την MDA υπήρξαν αυτές του **Von Stein** (1981) στη Γερμανία, του **Weibel**(1973) στην Ελβετία, των **Taffler** και **Tissaw**(1977) και του **Marais**(1979) στην Αγγλία, του **Bilderbeek** (1977) στην Ολλανδία και των **Altman** και **Lavallee** (1981) στον Καναδά. Σε όλες αυτές τις μελέτες, τα εκτιμημένα μοντέλα παρουσίασαν υψηλά ποσοστά επιτυχίας που κυμαίνονταν από 70% έως 90%. Παρόμοιες μελέτες του **Altman** (1973) στη Γαλλία και των **Castagna** και **Matolscy**(1981) στην Αυστραλία κατέληξαν σε μέτρια αποτελέσματα. Η Regression Analysis δεν βρήκε μεγάλη εφαρμογή σε χώρες εκτός των Ηνωμένων Πολιτειών.

Την ίδια περίοδο, έγιναν και κάποιες απόπειρες εφαρμογής λίγο διαφορετικών μοντέλων. Οι **Wilcox** (1971,1973,1976), οι **Santomero** και **Vinso** (1977) και ο **Vinso** (1979) βασίστηκαν σε μία διαφορετική θεωρία, την gambler's ruin theory, για τη δημιουργία μοντέλων πιστωτικού κινδύνου. Το gambler's ruin model της πιθανοτικής θεωρίας θεωρεί ότι ο παίκτης ξεκινά με ένα ορισμένο ποσό χρημάτων και κερδίζει ένα δολάριο με πιθανότητα p ενώ χάνει το ίδιο ποσό με πιθανότητα $1-p$. Το παιχνίδι συνεχίζεται μέχρι ο παίκτης ή ο αντίπαλος του να χάσει όλα του τα χρήματα. Η θεωρία που βασίζεται σε αυτό το απλό σενάριο είναι αρκετά καλά ανεπτυγμένη: έχουν εκφραστεί μαθηματικά η πιθανότητα της απόλυτης καταστροφής, του αναμενόμενου απόλυτου κέρδους ή ζημιάς του παίχτη, της αναμενόμενης διάρκειας του παιχνιδιού κτλ. Στις χρηματοοικονομικές εφαρμογές, η εταιρία θεωρείται ως παίκτης και η χρεωκοπία συμβαίνει όταν η καθαρή της αξία πέσει στο μηδέν. Οι απόπειρες παρολαυτά εφαρμογής αυτού του μοντέλου στο πρόβλημα του πιστωτικού κινδύνου υπήρξαν απογοητευτικές, ίσως γιατί η θεωρία ήταν πολύ απλή με το να υποθέτει ότι οι ταμειακές ροές της επιχείρησης προέρχονται από μία σειρά ανεξάρτητων δοκιμών, αγνοώντας την πιθανή επέμβαση της διοίκησης. Οι εφαρμογές των τριών παραπάνω αναλυτών, ήταν περισσότερο εμπειρικές χωρίς μάλιστα να υποστηρίζονται από στατιστικούς ελέγχους και τεστ.

3.7.ΑΝΑΛΥΤΙΚΗ ΠΑΡΟΥΣΙΑΣΗ ΕΜΠΕΙΡΙΚΩΝ ΜΕΛΕΤΩΝ

a) **Univariate approach / Beaver 1966**

Ο Beaver παρουσίασε την πρώτη μοντέρνα στατιστική εκτίμηση μοντέλων για την πρόβλεψη της χρηματοοικονομικής αποτυχίας. Από 30 δείκτες που εξέτασε ανακάλυψε ότι τρεις προβλέπουν καλύτερα την οικονομική αποτυχία: οι δείκτες ταμειακές ροές/σύνολο ενεργητικού, καθαρά κέρδη/ σύνολο υποχρεώσεων και ταμειακές ροές/ σύνολο υποχρεώσεων. Η έρευνα του Beaver αποδεικνύει ότι οι χρηματοοικονομικοί δείκτες από μόνοι τους μπορούν να προβλέψουν την αποτυχία αρκετά καλά. Η διαδικασία που ακολούθησε ο Beaver είχε ως εξής: για κάθε δεικτη βρήκε ένα cut-off point, ένα όριο, τέτοιο ώστε εταιρίες με δείκτες πάνω από αυτό το όριο να κατατάσσονται ως πιθανές μη αποτυχημένες, ενώ εταιρίες με μικρότερους δείκτες να κατατάσσονται ως πιθανές αποτυχημένες. Τα όρια αυτά προήλθαν από ένα αρχικό δείγμα, αλλά χρησιμοποιήθηκαν για να κατατάξουν εταιρίες σε ένα δευτερογενές δείγμα. Τα αποτελέσματα έδειξαν ότι ακόμη και 5 χρόνια πριν την αποτυχία, μόνο 22% των εταιριών τόσο του αρχικού όσο και του δευτερογενούς δείγματος, κατατάσσονται λάθος από τον δείκτη των ταμειακών ροών/ σύνολο υποχρεώσεων.

b) **MDA** :

Λίγα χρόνια αργότερα, άρχισε να γίνεται αισθητό από τους επόμενους αναλυτές ότι η στήριξη των μοντέλων πρόβλεψης χρεωκοπίας μόνο σε χρηματοοικονομικούς δείκτες παραήταν απλοποιημένη για να συλλάβει την πολυπλοκότητα της οικονομικής αποτυχίας. Αποτέλεσμα της παραπάνω παραδοχής ήταν η χρήση πιο πολύπλοκων μεθόδων που στηρίζονταν σε δυναμικούς συνδυασμούς χρηματοοικονομικών δεικτών, όπως ήταν η Multiple Discriminant Analysis (MDA).

i) **Altman Z-score 1968**

Ο Edward Altman υπήρξε ο πρώτος που προσπάθησε να γεφυρώσει την ανάλυση βάσει χρηματοοικονομικών δεικτών με τις καινούργιες στατιστικές τεχνικές, εφαρμόζοντας την MDA. Η διαδικασία που ακολούθησε ο Altman για να διαμορφώσει την MDA ήταν η εξής: καθόρισε τα groups (bankrupt-non bankrupt), συγκέντρωσε στοιχεία για τις εταιρίες που ανήκουν σε αυτά και σχημάτισε μια γραμμική συνάρτηση που με βάση τα παραπάνω στοιχεία – χαρακτηριστικά να διαχωρίζει όσο το δυνατόν καλύτερα τις παρατηρήσεις ανάμεσα στα 2 groups. Κατ'αυτόν τον τρόπο διαμόρφωσε μια γραμμική συνάρτηση παλινδρόμησης με διαχωριστικές ικανότητες, μια μορφή MDA, που ονομάστηκε z-score. Το z-score, ως ένας σύνθετος δείκτης οικονομικής αδυναμίας, βασίζεται σε αυτούς τους χρηματοοικονομικούς δείκτες των λογιστικών καταστάσεων που στο παρελθόν έχουν αποδειχτεί να έχουν τη μεγαλύτερη σχέση με τη μη βιωσιμότητα των επιχειρήσεων (corporate insolvency). Το z-score δείχνει τη συγκριτική θέση μιας εταιρίας ανάμεσα στις

αλλες οδηγώντας έτσι σε μια γρήγορη αναγνώριση των εταιριών που απαιτούν ιδιαίτερη προσοχή. Η γενική ιδέα του z-score βασίζεται σε έναν συνδυασμό χρηματοοικονομικών δεικτών ρευστότητας, κερδοφορίας, ταμειακών ροών, μόχλευσης και τιμής μετοχής σε έναν μοναδικό δείκτη που συνιστά το z-score. Αναλυτικότερα, η διαδικασία της μελέτης είχε ως εξής:

Δείγμα: επέλεξε 66 κατασκευαστικές εταιρίες από τις οποίες οι 33 χρεωκοπημένες και οι υπόλοιπες υγιείς με εύρος για το μέγεθος του ενεργητικού τους \$0,7-25,9 m.για τις πρώτες και \$1-25 m.για τις δεύτερες. Οι πολύ μεγάλες και πολύ μικρές εταιρίες αφαιρέθηκαν από το δείγμα.

Χρονική περίοδος: αφορά τις εταιρίες που χρεωκόπησαν την περίοδο 1946-1965 και οι ισολογισμοί ήταν 1 έτους πριν την κήρυξη σε χρεωκοπία.

Μεταβλητές: Αρχικά οι χρηματοοικονομικοί δείκτες ήταν 22, χωρισμένοι σε 5 κατηγορίες ρευστότητας, αποδοτικότητας, μόχλευσης, βιωσιμότητας και δραστηριότητας. Από τους 22 επιλέχθηκαν τελικά 5 που θεωρήθηκαν ότι συμβάλλουν περισσότερο στην ακριβή πρόβλεψη της χρεωκοπίας: x_1 = working capital/total assets, x_2 = retained earnings /total assets, x_3 = EBIT/total assets, x_4 = market value equity/book value of total debt, x_5 = sales/total assets

Μορφή MDA: $z = 0,012x_1 + 0,014x_2 + 0,033x_3 + 0,006x_4 + 0,999x_5$

Η παραπάνω εξίσωση είναι γνωστή και ως **Z-Altman score**.

Discriminating power of variables (τεστ για τη διακριτική ικανότητα των μεταβλητών):

i)με *F-test(univariate method)*: το F-test συσχετίζει τη διαφορά μεταξύ των μέσων τιμών των δεικτών σε κάθε group και των μεταβλητοτήτων τους, μας δείχνει δηλαδή πόσο διαφορετικές είναι κάθε φορά οι τιμές π.χ. της x_1 από τον μέσο της σε καθένα από τα δύο group ή αλλιώς πόσο επηρεάζει η x_1 την τιμή του z. Αν ασκεί μεγάλη επιρροή, τότε έχει discriminating power. Τα αποτελέσματα αυτού του τεστ έδειξαν ότι όλες οι μεταβλητές έχουν discriminating power εκτός από την x_5 η οποία όμως συμμετέχει στο μοντέλο λόγω της συσχέτισής της με τις άλλες όπως θα αποδειχτεί με το επόμενο τεστ.

ii)με *scaled vector(multivariate method)*: αποδεικνύει και αυτό την discriminating power όλων των μεταβλητών καθώς και της x_5 η οποία εμφανίζεται να έχει μια αρνητική συσχέτιση με την x_3 .

Discriminating power of model (τεστ για τη διακριτική ικανότητα του μοντέλου):

Με τη μεγιστοποίηση ενός λόγου, καταφέρνει να απομακρύνει τους μέσους των z-scores των 2 groups (best discriminate between groups) και να μειώσει την απόσταση μεταξύ των z-scores καθεμιάς εταιρίας και του μέσου z-score του group στο οποίο ανήκει η εταιρία (variables most similar within groups). Αυτό το τεστ ελέγχει δηλαδή κατά πόσον η MDA που κατασκεύασε διαχωρίζει όσο το δυνατόν καλύτερα ανάμεσα στα 2 groups και αυτό μπορεί να γίνει μόνο όταν η διακύμανση ανάμεσα στα δύο groups (variance between groups) είναι όσο το δυνατόν μεγαλύτερη σε σχέση με τη διακύμανση μέσα σε κάθε group (variance within groups), όπως είχαμε αναφέρει και στη θεωρητική ανάλυση της MDA.

Εμπειρικά Αποτελέσματα:

i) Αρχικό δείγμα: το αρχικό δείγμα των 33 εταιριών κάθε group εξετάστηκε με βάση στοιχεία ισολογισμών 1 έτους προ της χρεωκοπίας. Από τη στιγμή που οι

discriminant coefficients και οι κατανομές των groups προήλθαν από αυτό το δείγμα, αναμένεται υψηλός βαθμός ακρίβειας κατάταξης. Το μοντέλο αποδείχτηκε πράγματι εξαιρετικά ακριβές κατατάσσοντας σωστά το 95% του συνολικού δείγματος. Το Type I error (πρόβλεψη μη χρεωκοπίας για εταιρία που τελικά χρεωκοπεί) ήταν μόνο 6%, ενώ το Type II error (πρόβλεψη χρεωκοπίας για εταιρία που τελικά δεν χρεωκοπεί) ήταν 3%. Υπάρχει επομένως μια upward bias, με την έννοια ότι το μοντέλο κατατάσσει ως χρεωκοπημένες περισσότερες εταιρίες από ό,τι θα έπρεπε.

ii) Αποτελέσματα για δεδομένα 2 ετών προ χρεωκοπίας: όπως ήταν αναμενόμενο, η ακρίβεια του μοντέλου μειώθηκε στο 72%.

iii) Εξέταση bias: η πιθανή bias οφείλεται στη μείωση από 22 σε 5 μεταβλητές. Μπορεί ένα υποσύνολο μεταβλητών να είναι αποτελεσματικό για το αρχικό δείγμα αλλά όχι και για ολόκληρον τον πληθυσμό. Εξέτασε λοιπόν αν υπάρχει τέτοιο bias χρησιμοποιώντας ένα υποσύνολο του αρχικού δείγματος για να υπολογίσει από την αρχή τις παραμέτρους του μοντέλου, και με το υπόλοιπο του αρχικού δείγματος κατέταξε τις παρατηρήσεις βάσει των νέων παραμέτρων που υπολόγισε με το υπο-δείγμα. Στη συνέχεια εφαρμοσε ένα t-test για τον έλεγχο της σημαντικότητας των αποτελεσμάτων της παραπάνω διαδικασίας. Τα αποτελέσματα έδειξαν ότι το μοντέλο διαθέτει discriminating power και σε παρατηρήσεις άλλες από αυτές που χρησιμοποιήθηκαν για την εκτίμηση των παραμέτρων.

iv) Τεστ με δευτερογενές δείγμα χρεωκοπημένων εταιριών: Επέλεξε άλλο δείγμα 25 χρεωκοπημένων εταιριών παρόμοιες με τις προηγούμενες του αρχικού δείγματος. Χρησιμοποίησε τις ίδιες παραμέτρους του αρχικού δείγματος και τα αποτελέσματα έδειξαν παραδόξως ακόμη μεγαλύτερη ακρίβεια από του αρχικού δείγματος.

v) Τεστ με δευτερογενές δείγμα υγιών εταιριών: επέλεξε άλλο δείγμα 66 εταιριών όχι χρεωκοπημένων ούτε υγιών, αλλά με οικονομικά προβλήματα για 2 έτη. Τα αποτελέσματα έδειξαν ακρίβεια 79% αλλά με το 11% των εταιριών με οικονομικά προβλήματα να βρίσκεται στην gray area του z.

vi) Τεστ για μεγαλύτερη χρονική περίοδο: Χρησιμοποίησε δεδομένα για 5 χρόνια στο δείγμα των αρχικών 33 χρεωκοπημένων εταιριών και διαπίστωσε ότι όσο αυξάνεται ο χρόνος τόσο μειώνεται η ακριβεια του μοντέλου.

Γενικά συμπεράσματα: η MDA δίνει ακριβή πρόβλεψη για την αποτυχία των εταιριών μέχρι και 2 χρόνια πριν από αυτήν. Αν $z < 1,81$ η εταιρία κρίνεται ότι θα χρεωκοπήσει, ενώ αν $z > 2,99$ η εταιρία κατατάσσεται στις υγιείς. Η περιοχή 1,81-2,99 κρίνεται ως gray area όπου δεν γνωρίζουμε τί ακριβώς συμβαίνει.

ii) Zeta model 1977

Εννέα χρόνια αργότερα, ο Altman σε συνεργασία με δύο άλλους αναλυτές, τον R.G.Haldeman και τον P.Narayanan, θεώρησαν αναγκαία τη βελτίωση του παραπάνω μοντέλου και την προσαρμογή του στα σύγχρονα μεταβαλλόμενα δεδομένα. Σημαντικές αλλαγές της αγοράς και της επιχειρησιακής ζωής, όπως η χρεωκοπία ολοένα και μεγαλύτερων εταιριών, η αλλαγή των λογιστικών προτύπων και η αύξουσα σημασία άλλων κλάδων εταιριών εκτός των κατασκευαστικών, έπρεπε να ενταχθούν στο μοντέλο. Το νέο μοντέλο που δημιουργήθηκε γνωστό με το όνομα ZETA analysis και που δεν είναι παρά μια βελτιωμένη έκδοση του προηγούμενου z-score, λαμβάνει υπόψη τις παραπάνω εξελίξεις και μελετάει βαθύτερα τις αμφιλεγόμενες όψεις της discriminant analysis και την αναπόφευκτη χρονική υστέρηση της πρόβλεψης. Η διαδικασία που ακολουθήθηκε είχε ως εξής:

Δείγμα: περιείχε 53 χρεωκοπημένες και 58 υγιείς εταιρίες αντιστοιχισμένες κατά κλάδο και χρονική περίοδο δεδομένων. Το δείγμα ήταν περίπου ίσα μοιρασμένο ανάμεσα σε κατασκευαστικές και εμπορικές εταιρίες και το μέσο μέγεθος ενεργητικού των χρεωκοπημένων εταιριών κυμαινόταν στα \$100m.

Χρονική περίοδος: το 94% του συνόλου των χρεωκοπημένων εταιριών χρεωκόπησαν στην περίοδο 1969-1975.

Μεταβλητές: επέλεξαν 27 δείκτες που διακρίνονταν σε πέντε κατηγορίες: αποδοτικότητα, ρευστότητας, μόχλευσης, κεφαλαιοποίησης και μεταβλητότητας κερδών. Μετά από μια iterative process μείωσης του αριθμού των μεταβλητών, κατεληξαν σε ένα μοντέλο με επτά μεταβλητές. Περισσότερες ή λιγότερες από αυτές, αποδείχτηκε ότι δεν θα πρόσφεραν τίποτα παραπάνω στην ακρίβεια του μοντέλου. Οι 7 μεταβλητές ήταν οι εξής: x_1 = return on assets= EBIT/ total assets, x_2 = stability of earnings, x_3 =debt service= EBIT/ total interest payments, x_4 = cumulative profitability= retained earnings/ total assets, x_5 = liquidity, x_6 = capitalization= equity/ total capital, x_7 = size.

Λογιστικές προσαρμογές: οι αναλυτές προσάρμοσαν τα λογιστικά δεδομένα στις πρόσφατες και σημαντικότερες λογιστικές εξελίξεις.

Στατιστική μεθοδολογία: χρησιμοποίησαν το τεστ H_1 για τον προσδιορισμό της κατάλληλης μεθόδου ανάμεσα στη γραμμική(linear) και την τετραγωνική(quadratic). Σύμφωνα με αυτό το τεστ, όταν οι πίνακες συνδιακυμάνσεων των G groups είναι στατιστικά όμοιοι τότε η γραμμική δομή είναι κατάλληλη για την ανάλυση των χαρακτηριστικών κατάταξης ενός συγκεκριμένου δείγματος. Αντίθετα, όταν οι πίνακες συνδιακυμάνσεων είναι ανόμοιοι, η τετραγωνική δομή είναι η κατάλληλη. Το H_1 test έδειξε ότι οι group dispersion matrices είναι not identical. Επομένως κατάλληλη είναι η quadratic form. Παρολαυτά, για το original sample(αρχικό δείγμα), τα αποτελέσματα έδειξαν ότι και η linear και η quadratic είχαν την ίδια ακρίβεια. Τα αποτελέσματα όμως των holdout samples(δευτερογενών δειγμάτων) έδειξαν σαφή ανωτερότητα της γραμμικής δομής. Υπήρξε λοιπόν ένα δίλημμα ως προς ποια δομή τελικά έπρεπε να ακολουθήσουν, μέχρι που αποφασίστηκε η χρήση της γραμμικής καθώς όλα τα σχετικά τεστ που διεξήχθησαν αργότερα βασίζονταν στο γραμμικό μοντέλο και καθώς έτσι θα μπορούσε να

αποφευχθεί η πιθανή υψηλή ευαισθησία που θα εμφάνιζαν οι τετραγωνικές παράμετροι στις ατομικές παρατηρήσεις του δείγματος.

i) Discriminating power of variables (τεστ για τη διακριτική ικανότητα των μεταβλητών): η σχετική σημαντικότητα της καθεμιάς μεταβλητής στον καθορισμό του z ελέγχεται με 6 διαφορετικά τεστ, τα αποτελέσματα των οποίων συγκλίνουν σχεδόν απόλυτα κατατάσσοντας τις μεταβλητές κατά σειρά σημαντικότητας ως εξής: $x_4, x_2, x_6, x_7, x_5, x_3, x_1$.

ii) Discriminating power of model (tests of model validity ή τεστ για τη διακριτική ικανότητα του μοντέλου): έγιναν τεστ στο αρχικό δείγμα σε δεδομένα 1 έτους πριν τη χρεωκοπία, Lachenbruch validation tests και τεστ σε δευτερογενή δείγματα για 2-5 χρόνια πριν τη χρεωκοπία. Τα αποτελέσματα των τεστ αναφέρονται και στη γραμμική και στην τετραγωνική δομή του 7-μεταβλητού μοντέλου. Η ακρίβεια κατάταξης για το χρεωκοπημένο group μειώνεται όσο τα δεδομένα απομακρύνονται χρονικά από τη στιγμή της χρεωκοπίας. Συγκεκριμένα, παρατηρείται ακρίβεια 70% για 5 χρόνια προ της χρεωκοπίας. Η ακρίβεια παρουσιάζεται ίδια και για τις δυο μορφές δομής του μοντέλου στο τεστ με το αρχικό δείγμα. Στα δευτερογενή δείγματα όμως, η γραμμική εμφανίζεται ακριβέστερη. Ένα τελευταίο τεστ ακρίβειας έγινε με βάση την τυχαία επιλογή των μισών παρατηρήσεων έτσι ώστε να επιβεβαιωθούν οι παράμετροι του μοντέλου με τη χρήση των υπόλοιπων παρατηρήσεων ως δευτερογενή δείγματα. Η ακρίβεια και με αυτό το τεστ κυμαίνεται και πάλι σε εντυπωσιακά επίπεδα ακόμη και για τα 5 χρόνια της ανάλυσης, μειούμενη βέβαια καθώς προχωράμε πίσω στο χρόνο.

Σύγκριση με z-score: Τρόποι σύγκρισης :

-Απλή σύγκριση των 2 μοντέλων με τα αρχικά δείγματα του καθενός για 5 χρόνια προ χρεωκοπίας → ίδια ακρίβεια στον 1^ο χρόνο προ χρεωκοπίας, ZETA πιο ακριβές στα χρόνια 2-5 στην bankruptcy classification, z-score πιο ακριβές στα ίδια χρόνια στην non- bankruptcy classification.

-Εφαρμογή του παλιού μοντέλου στο νέο δείγμα του ZETA → ZETA πιο ακριβές.

-Με βάση τις 5 μεταβλητές του z-score και το νέο δείγμα του ZETA υπολογισμός εκ νέου των παραμέτρων → ZETA πιο ακριβές.

-Σύγκριση F-statistic univariate tests για τις 5 μεταβλητές → ZETA πιο ακριβές.

Group prior probabilities, error costs, model efficiency: Αν θέσουμε ίσες τις prior probabilities να χρεωκοπήσει ή όχι μια εταιρία, το γραμμικό μοντέλο θα καταλήξει σε ένα cutoff ή critical score = 0.

Αν η εταιρία έχει score >0 → non-bankrupt

Αν η εταιρία έχει score <0 → bankrupt

Στο ίδιο μηδενικό cutoff score θα καταλήξουμε και αν ελαχιστοποιήσουμε το συνολικό κόστος της misclassification :

optimal cutoff score $ZETA = ZETA_c = \ln(q_1 c_I / q_2 c_{II})$.

Για να συγκρίνουμε την efficiency του ZETA με άλλων μοντέλων βρίσκουμε το αναμενόμενο κόστος του: $EC_{ZETA} = q_1 (M_{12}/N_1) C_I + q_2 (M_{12}/N_2) C_{II}$.

Το ότι θεωρήσαμε όμως ίσες τις prior probabilities και τα costs of errors μπορεί να μας οδηγήσει σε bias.

Cost of classification errors: Μία εναλλακτική προσέγγιση του zero cutoff score είναι το cost of errors in classification. Ορίζονται ως:

Type I classification error: αποδοχή δανείου που δεν θα αποπληρωθεί

Type II classification error: απόρριψη δανείου που θα αποπληρωθεί

C_I : cost of Type I error: $C_I = 1 - LLR/GLL$

C_{II} : cost of Type II error: $C_{II} = r - I$

Revised cutoff score and model efficiency tests:

Ύστερα από υπολογισμούς, το cutoff score του ZETA υπολογίστηκε σε : $ZETA_c = -0,337$

Η ακρίβεια του ZETA θα μεταβληθεί με ένα νέο cutoff score. Η σύγκριση που έγινε μεταξύ του ZETA και δύο άλλων μεθόδων με βάση τα costs of errors έδειξαν σαφή ανωτερότητα του ZETA ως προς την αποτελεσματικότητά του.

Γενικά συμπεράσματα: Το νέο μοντέλο ZETA αποδείχτηκε ακριβές για μέχρι και 5 έτη πριν τη χρεωκοπία με επιτυχία 90% για τον 1^ο χρόνο και 70% για 5 χρόνια πριν. Περίεργο ήταν ότι παρόλο που οι στατιστικές ιδιότητες των στοιχείων υποδείκνυαν τη χρήση quadratic structure, η linear structure του ίδιου μοντέλου είναι καλύτερη από την quadratic σε tests of model validity. Το ZETA αποδείχτηκε καλύτερο σε σύγκριση με άλλες τεχνικές.

iii) Altman Z-score / Bank of England 1982

Το 1982 η Κεντρική Τράπεζα της Αγγλίας εφάρμοσε το z-score σε ένα δείγμα αγγλικών βιομηχανικών επιχειρήσεων.

Δείγμα: το δείγμα αποτελούταν από 38 αποτυχημένες και 53 υγιείς βιομηχανικές εταιρίες.

Χρονική περίοδος: δεδομένα μέχρι και 3 ετών πριν την αποτυχία.

Εμπειρικά αποτελέσματα:

Έλεγχος της ακρίβειας κατάταξης του μοντέλου (tests of model validity): το μοντέλο βρέθηκε περισσότερο ακριβές μέχρι και 3 έτη πριν την αποτυχία. Ένα τεστ εγκυρότητας όμως που διεξήχθη σε δείγμα διαφορετικό από το αρχικό με 10 αποτυχημένες και 19 υγιείς εταιρίες, έδειξε misclassification πάνω από 50% των υγιών εταιριών. Παρολαυτά, το νέο αυτό δείγμα περιείχε σκόπιμα έναν αριθμό εταιριών που αντιμετώπιζαν οικονομικά προβλήματα, οι περισσότερες από τις οποίες αντιστοιχήθηκαν σε χαμηλά z-scores, γεγονός που επισήμανε κίνδυνο αποτυχίας και μετρίασε την παραπάνω misclassification. Το 50% των εταιριών λοιπόν κρίθηκε επικίνδυνο με z-score <0,7 ή με κατάταξη τους στην gray area και για παραπάνω από 1 χρόνο, η ακρίβεια του μοντέλου μειώθηκε σε συνάρτηση με τον χρόνο.

Γενικά συμπεράσματα: Η τράπεζα της Αγγλίας σύμφωνα με την παραπάνω έρευνα, θεώρησε ότι το z-score πρέπει να χρησιμοποιείται μόνο ως οδηγός για την αναγνώριση και επισήμανση των προβληματικών εταιριών. Δεν το αποδέχτηκε ως καταλυτικό μέτρο πρόβλεψης για τη χρεωκοπία μιας εταιρίας. Γι'αυτήν, δεν αποτελεί παρά μόνο το πρώτο βήμα για περαιτέρω εξέταση κάποιων εταιριών.

Επισημαίνει επίσης ότι: το z-score πρέπει να χρησιμοποιείται με προσοχή γιατί: οι λογιστικές διαδικασίες και τα προφίλ των εταιριών διαφέρουν κάνοντας δύσκολη τη σύγκριση και τα δεδομένα λαμβάνονται με μια χρονική υστέρηση.

Γενικά θεωρεί ότι το z-score πρέπει να χρησιμοποιείται μαζί με άλλες μεθόδους.

iv) Joint discriminant functions and neural networks 1994

Το 1994 η Κεντρική Τράπεζα της Ιταλίας δημιούργησε ένα σύστημα γρήγορης ανίχνευσης προβληματικών εταιριών. Το σύστημα αυτό βασίστηκε αρχικά σε ένα ζεύγος linear discriminant functions που δούλευαν παράλληλα και ήταν προσαρμοσμένες σε βιομηχανικές εταιρίες. Στη συνέχεια, η τράπεζα βελτίωσε αυτό το σύστημα γρήγορης ανίχνευσης με τη χρήση των προηγμένων τεχνολογικά neural networks.

Δείγμα: το δείγμα που χρησιμοποιήθηκε και στις δύο μεθόδους περιείχε 213 μη υγιείς και 213 υγιείς μικρομεσαίες επιχειρήσεις.

Χρονική περίοδος: το δείγμα αφορούσε εταιρίες που απέτυχαν μέσα στην περίοδο 1985-1992 και τα στοιχεία συλλέχτηκαν για 3 έτη πριν την αποτυχία.

Στατιστικές μέθοδοι:

A) Joint Discriminant Functions (JDF):

Δομή: Η discriminant analysis παρουσιάζει μειωμένη ακρίβεια σε μεγάλο δείγμα λόγω των πολλών περιπτώσεων που εισέρχονται. Δεν έχει δηλαδή την ικανότητα να συγκρίνει ανάμεσα σε εταιρίες outstanding, normal και vulnerable παρά μόνο ανάμεσα σε healthy και unsound. Για την αντιμετώπιση του παραπάνω προβλήματος, χωρίσαν τη διαδικασία σε δύο στάδια, σε δύο discriminant functions: η F_1 : διακρίνει μεταξύ healthy και unsound

η F_2 : διακρίνει μεταξύ vulnerable και unsound

Η F_2 έρχεται σε εφαρμογή μετά την F_1 , δηλαδή αφότου γίνει η διάκριση ανάμεσα σε υγιείς και μη εταιρίες ακολουθεί περαιτέρω διάκριση των από την προηγούμενη διαδικασία καταταγμένων εταιριών ως μη υγιείς, σε ευαίσθητες και μη υγιείς.

Εμπειρικά αποτελέσματα: Οι δείκτες για την διαμόρφωση των F_1 , F_2 στηρίχθηκαν σε οικονομικά στοιχεία του 3^{ου} χρόνου πριν την αποτυχία. Η διάγνωση επομένως έγινε με βάση μια joint analysis των 2 functions F_1 , F_2 , με την F_2 να παρουσιάζει μικρότερη discriminating capacity, ειδικά για τις unsound εταιρίες. Η discriminant analysis παρουσίασε ακρίβεια 90,3% για τις υγιείς εταιρίες και 86,4% για τις προβληματικές.

B) Neural Networks (NN)

Δομή: Τα Neural Networks είναι μια τεχνολογική εξέλιξη της πληροφορικής που στηρίζεται στη μίμηση της διαδικασίας σκέψης και λήψης αποφάσεων του ανθρώπινου εγκεφάλου. Τα νευρώνια (neurons) αποτελούν την μικρότερη δομική μονάδα σκέψης του ανθρώπινου μυαλού. Τα Neural Networks αποτελούνται από έναν μεγάλο αριθμό στοιχειωδών processing units που ονομάζονται νευρώνια, από τα οποία το καθένα είναι ικανό να διεξάγει σχετικά

απλούς υπολογισμούς και είναι ταυτόχρονα συνδεδεμένο με όλα τα υπόλοιπα. Το αποτέλεσμα προκύπτει από τη συλλογική τους συμπεριφορά παρά από την ιδιαίτερη συμπεριφορά του καθενός. Οι σύνδεσμοι δεν είναι αυστηροί αλλά μπορούν να μετατραπούν μέσω κάποιων διαδικασιών εκμάθησης (learning processes) που γίνονται από την interaction του network με τον εξωτερικό κόσμο ή με ένα set συμβολικών σημάτων. Κάθε unit i δέχεται ένα input x_i απ' έξω ή από άλλα neurons με τα οποία συνδέεται με ένα weight w_{ji} . Το output y_i μπορεί να είναι μια ολική ανταπόκριση του network (final output value) ή ένα νέο input για άλλα neuron units.

Πολυπλοκότητα: Η πολυπλοκότητα των networks εμφανίζεται σε διάφορα επίπεδα, single neuron layer που επιτρέπει γραμμικούς διαχωρισμούς του χώρου δεδομένων, two layers που επιτρέπουν πολύπλοκα γεωμετρικά σχήματα, three layers που επιτρέπουν σχήματα οποιασδήποτε μορφής. Δεν υπάρχουν κανόνες για τον ιδανικό βαθμό πολυπλοκότητας ενός network.

Learning cycles: Το network λοιπόν είναι ένα δεδομένο set of inputs που 'γεννούν' μια response η οποία συγκρίνεται με την απαιτούμενη response. Αν η given response \neq required response τα weightings αλλάζουν μέσω μιας διαδικασίας εκμάθησης (learning mechanism/cycle). Όταν το λάθος μειωθεί σε ένα αποδεκτό επίπεδο η εκμάθηση σταματάει και τα weightings 'κλειδώνονται'.

Πλεονεκτήματα: Τα πλεονεκτήματα του network είναι ότι δεν απαιτεί προσδιορισμό από πριν των λειτουργικών μορφών, δεν υιοθετεί περιοριστικές υποθέσεις για τα χαρακτηριστικά των στατιστικών κατανομών των μεταβλητών και των λαθών, δουλεύει με ανακριβείς μεταβλητές και είναι δυναμικό δηλαδή προσαρμόζεται.

Μειονεκτήματα: Το βασικό του μειονέκτημα είναι ότι δεν χαρακτηρίζεται από διαφάνεια, δεν μπορούμε με άλλα λόγια να δούμε την επιρροή των ενδογενών μεταβλητών στα neurons και το βαθμό συσχέτισης τους με καθένα από αυτά. Αυτό κάνει δύσκολη την αναγνώριση των individual contributions των inputs και τη διαμόρφωση του output πολύπλοκη και αβέβαιη. Επίσης η συμπεριφορά του network μπορεί να είναι περίεργη, απρόβλεπτη και αντίθετη στη λογική, αν μεταβάλλουμε για παράδειγμα έναν μόνο δείκτη μπορεί η εταιρία από εκεί που πριν είχε κριθεί ως υγιής να θεωρηθεί τώρα αποτυχημένη.

Εμπειρικά αποτελέσματα: Το πειραματικό πρόγραμμα των neural networks που ακολούθησε η τράπεζα της Ιταλίας αναπτύχθηκε στα παρακάτω βήματα:

- έλεγχος της ικανότητας του network να αναπαράγει τις αριθμητικές αξίες των scores της linear discriminant analysis, με inputs όμως δείκτες διαφορετικούς από αυτούς της discriminant analysis \rightarrow το network approximates
- έλεγχος της ικανότητας του να διακρίνει τα δείγματα μεταξύ bankrupt και healthy εταιρίες \rightarrow επιτυχής
- ένταξη της μεταβολής της απόδοσης της εταιρίας μέσα στον χρόνο. Αν το profile του 2^{ου} χρόνου ακολουθεί μια τάση που συμφωνεί με την τάση της προηγούμενης περιόδου τα συμπεράσματα ενισχύονται, αν όχι, αποδυναμώνονται.

- Έλεγχος της ικανότητας του να διαχωρίζει ανάμεσα σε healthy, vulnerable, unsound. Στην περίπτωση αυτή χρησιμοποίησαν 2 output neurons, το πρώτο για τον διαχωρισμό μεταξύ healthy, unsound και το δεύτερο για τη διάκριση μεταξύ unsound και vulnerable. (όπως χρησιμοποίησαν δύο discriminant functions F_1 , F_2 στην περίπτωση της Joint Discriminant Functions μεθόδου.)

Τα καλύτερα αποτελέσματα έδωσε το 3-layer network με χρήση 15 δεικτών ρευστότητας, αποδοτικότητας, εσωτερικής χρηματοδότησης και χρηματοοικονομικής δομής. Αποδείχτηκε ότι αν η φάση εκμάθησης ήταν μεγαλύτερη το λάθος θα μειωνόταν. Το 3-layer network με 15 δείκτες και 2000 κύκλους εκμάθησης εμφάνισε ακρίβεια 97,7% για τις unsound εταιρίες και 97% για τις healthy. Αποδείχτηκε το καλύτερο στην κατηγορία του και καλύτερο και από την discriminant analysis.

Σύγκριση NN και JDF :

Χρησιμοποίησαν 9 από τους 11 δείκτες της F_1 discriminant function στα networks. Τα αποτελέσματα αποδείχτηκαν χαμηλότερα ή σχεδόν όμοια με της discriminant function. Παρολαυτά δεν είναι σίγουρο ότι η μορφοποίηση της γνώσης μέσα στο network είναι απόλυτα ισότιμη με αυτήν της discriminant function. Η discriminant function συμπεριφέρεται το ίδιο όταν οι τιμές των εξωγενών μεταβλητών ποικίλλουν, σε αντίθεση με το network.

Γενικά συμπεράσματα: Τα αποτελέσματα του network αποδείχτηκαν όμοια ή καλύτερα από της discriminant function εκτός από αυτά της control period ή των holdout samples που αποδείχτηκαν χειρότερα. Το network παρουσιάζει παράλογη συμπεριφορά που αυξάνεται με την πολυπλοκότητά του, δεν μας δίνει την δυνατότητα να ανιχνεύσουμε λάθη στις μεταβλητές ώστε να τα διορθώσουμε ούτε την ευκαιρία να εκτιμήσουμε την soundness των μεταβλητών a priori. Η discriminant function από την άλλη, διακρίνεται από μεγαλύτερη ταχύτητα εκτίμησης, χαμηλότερο κόστος και συνεπή συμπεριφορά για κάθε μεταβλητή. Μας δίνει επίσης τη δυνατότητα να ερμηνεύσουμε τη λειτουργική λογική του μοντέλου με βάση τις παραμέτρους και να γνωρίζουμε ανά πάσα στιγμή ποιές είναι οι σημαντικές μεταβλητές. Με βάση λοιπόν τα ιδιαίτερα χαρακτηριστικά των δύο μεθόδων, οι ερευνητές προτείνουν την παράλληλη χρήση και συνδυασμό και των δύο.

c) LPM / Linear Probability Model and Discriminant Analysis

εφαρμογή στην Ελλάδα 1984 :

Το 1988 ο Γ.Γκλούμπος και ο Θ.Γραμματικός προσπάθησαν να ελέγξουν και να αξιολογήσουν μέσω μιας μελέτης την προβλεπτική ικανότητα των μοντέλων πιστωτικού κινδύνου σε ελληνικά δεδομένα.

Δείγμα: το δείγμα που χρησιμοποίησαν περιείχε 29 χρεωκοπημένες και 29 αντιστοιχισμένες υγιείς βιομηχανικές ελληνικές εταιρίες. Η αντιστοίχιση έγινε με βάση τον κλάδο, το μέγεθος ενεργητικού και τη χρονική περίοδο.

Χρονική περίοδος: η χρονική περίοδος που τους ενδιέφερε ήταν τα έτη 1977-1981, στη διάρκεια των οποίων οι παραπάνω εταιρίες χρεωκόπησαν ή

δήλωσαν πτώχευση. Οι ισολογισμοί που χρησιμοποιήθηκαν ήταν μέχρι 3 ετών πριν τη χρεωκοπία.

Λογιστικές προσαρμογές: πριν τον υπολογισμό των χρηματοοικονομικών δεικτών έγιναν κάποιες προσαρμογές στις λογιστικές καταστάσεις ώστε να επιτευχθεί ομοιομορφία για όλες τις επιχειρήσεις.

Μεταβλητές: χρησιμοποιήθηκαν 17 λογιστικοί δείκτες, ένας αριθμός μεγαλύτερος από τον συνηθισμένο των προηγούμενων ξένων μελετών, με την αιτιολογία της εξαγωγής όσο το δυνατόν περισσότερων πληροφοριών.

Στατιστική μεθοδολογία: οι στατιστικές μέθοδοι που χρησιμοποιήθηκαν ήταν η Multiple Discriminant Analysis και το Linear Probability Model.

Εμπειρικά αποτελέσματα: για την επιλογή των καλύτερων discriminating μεταβλητών, εφαρμόστηκε ένα κριτήριο επιλογής για καθμία από τις δύο μεθόδους. Για την MDA εφαρμόστηκε το Wilk's criterion, που είναι το συνολικό πολυμεταβλητό F-ratio για τον έλεγχο των διαφορών ανάμεσα στα group centroids. Η μεταβλητή που μεγιστοποιεί το F-ratio ελαχιστοποιεί ταυτόχρονα και το λάμδα του Wilk, που είναι ένα μέτρο διαχωρισμού των groups (measure of group discrimination). Για το LPM χρησιμοποιήθηκαν μόνο εκείνες οι ανεξαρτητες μεταβλητές που κάλυπταν συγκεκριμένα στατιστικά κριτήρια. Η επιλογή των μεταβλητών έγινε με βάση τη συμμετοχή τους στην explained variance.

Τα μοντέλα εκτιμήθηκαν με τη βοήθεια του στατιστικού πακέτου SPSS. Η τελική μορφή των μοντέλων με βάση λογιστικά στοιχεία ενός έτους πριν τη χρεωκοπία, ήταν η εξής:

$$\text{MDA: } z = -0,863 - 2,461x_1 + 5,330x_2 - 0,022x_3 + 3,676x_4 + 3,543x_5 + 4,223x_6$$

$$\text{LPM: } y = 0,313 + 0,546x_2 + 805x_5 + 0,979x_6$$

Όπου : x_1 = current assets / total assets

x_2 = net working capital / total assets

x_3 = inventories / net working capital

x_4 = notes payable / total assets

x_5 = earnings after taxes / current liabilities

x_6 = gross income / total assets

z = overall z-score

y = overall y-score

Τα εκτιμημένα πρόσημα των παραπάνω μεταβλητών συμφωνούν με την υπόθεση τους εκτός από των x_1 , x_4 της MDA. Μια μεγαλύτερη μοχλευση και ένας μικρότερος δείκτης current assets / total assets αυξάνουν την πιθανότητα επιβίωσης, κάτι που είναι βέβαια αντίθετο της λογικής. Η απάντηση σε αυτό το παράδοξο σύμφωνα με τους αναλυτές βρίσκεται στην ουσία και την ιδιαιτερότητα του μοντέλου MDA. Δεν μπορούμε να εκτιμήσουμε τη σχετική σημαντικότητα μιας μεταβλητής ceteris paribus. Καθεμία ασκεί μια επιρροή στο z σε συνεργασία με όλες τις υπόλοιπες και όχι μόνη της. Το LPM αντίθετα, δεν παρουσίασε προβλήματα ερμηνείας. Σημειωτέον ότι και οι τρεις σημαντικές μεταβλητές του LPM είναι σημαντικές και για την MDA. Επομένως οι μεταβλητές x_2 , x_5 , x_6 αποτελούν τις σημαντικότερες discriminating variables.

Discriminating power of variables (τεστ για τη διακριτική ικανότητα των μεταβλητών):

i) F-test → στατιστικά σημαντικότερες μεταβλητές: x_2, x_5, x_6

ii) standardized coefficients test (multivariate method) → στατιστικά σημαντικότερες μεταβλητές: x_2, x_4, x_5, x_6 . Παρατηρούμε την προσθήκη της μεταβλητής x_4 , ακριβώς γιατί αυτό το τεστ αποτελεί μια πολυμεταβλητή μέθοδο που λαμβάνει υπόψη τη συνεργασία κάθε μεταβλητής με όλες τις άλλες, κάτι που δεν γινόταν με το F-test.

Discriminating power of model (τεστ για τη διακριτική ικανότητα του μοντέλου):

Με μέτρο την canonical correlation για την MDA και το multiple R για το LPM, καθώς και με ένα F-test για το καθένα, καταλήγουν σε υψηλή discriminating power και για τα δύο μοντέλα.

Εμπειρικά αποτελέσματα:(classification performance):

1 έτος προ χρεωκοπίας: ακρίβεια 91,4% και για τα 2 μοντέλα.

2 έτη προ χρεωκοπίας: ακρίβεια 78% για MDA και 76% για LPM. Μειώνεται όπως ήταν αναμενόμενο η ακρίβεια και αυξάνεται το type I error, που σημαίνει ότι υπάρχει μεγαλύτερη πιθανότητα λάθους κατάταξης για τις χρεωκοπημένες εταιρίες. Αυτό μειώνει κάπως την πρακτική χρησιμότητα των μοντέλων, καθώς προτιμούμε πάνω απ' όλα να προβλέπουμε τη χρεωκοπία παρά τη μη χρεωκοπία. Παρολαυτά, τα μοντέλα εξακολουθούν να παρουσιάζονται επιτυχή με ακρίβεια κατάταξης των χρεωκοπημένων εταιριών 60% για την MDA και 68% για το LPM.

3 έτη προ χρεωκοπίας: η ακρίβεια μειώνεται στο 70% για την MDA και αυξάνεται στο 78% για το LPM. Το LPM παρουσιάζεται σαφώς καλύτερο της MDA και στην συνολική ακρίβεια και στο type I error.

Τεστ για την robustness of models: με ένα validation test που χρησιμοποίησε ένα subset του αρχικού δείγματος για την επανεκτίμηση των παραμέτρων και την κατάταξη του υπόλοιπου αρχικού δείγματος σύμφωνα με τις νέες παραμέτρους και με ένα t-test για την στατιστική σημαντικότητα των παραπάνω αποτελεσμάτων, βρήκαν ότι τα μοντέλα διαθέτουν discriminating power και για παρατηρήσεις άλλες από τις αρχικές που χρησιμοποιήθηκαν για την εκτίμηση των παραμέτρων καθώς και ότι η MDA χαρακτηρίζεται από type I error bias.

Γενικά συμπεράσματα: Η MDA αποδείχτηκε πιο σύνθετη και εξελιγμένη αλλά με προβλήματα ερμηνείας, ενώ το LPM είναι πιο robust, προσφέρει καλύτερες προβλέψεις για 2-3 χρόνια πριν τη χρεωκοπία και είναι γενικά περισσότερο αποτελεσματικό λόγω της μεγαλύτερης ακρίβειάς του στην κατάταξη των χρεωκοπημένων εταιριών.

d) Logit /Logit Regression approach / Martin 1977:

Το 1977, ο Daniel Martin εφάρμοσε για πρώτη φορά το logit model στο πρόβλημα έγκαιρης προειδοποίησης για την αποτυχία τραπεζικών ιδρυμάτων (bank early warning problem). Βασικός σκοπός της έρευνάς του υπήρξε η κατασκευή ενός μοντέλου έγκαιρης προειδοποίησης που να εκφράζει την πιθανότητα μελλοντικής αποτυχίας τραπεζών ως συνάρτηση μεταβλητών

προερχομένων από λογιστικές καταστάσεις τρέχουσας περιόδου. Τυπικά, το μοντέλο εκτιμήθηκε σε παρελθοντικά οικονομικά στοιχεία. Η εμπειρική προσέγγιση εξαρτάται από την αναγνώριση των τραπεζών που απέτυχαν στο παρελθόν. Η ομαδοποίηση αυτή των τραπεζών αναλύεται σε όρους ανεξάρτητων μεταβλητών διαφορετικής παρελθούσας χρονικής περιόδου. Το αποτέλεσμα είναι μια συνάρτηση που συσχετίζει την πιθανότητα αποτυχίας της περιόδου t με τις ανεξάρτητες μεταβλητές της περιόδου $t-1$. Όπως ισχύει για κάθε άλλο μοντέλο, η χρησιμότητα του εξαρτάται από τη σταθερότητά του μέσα στο χρόνο και την ποιότητα των δεδομένων. Για τη δημιουργία του μοντέλου logit επομένως, ο Martin υπέθεσε ότι η ανεξάρτητη μεταβλητή x_i ανήκει σε μια χρονική περίοδο νωρίτερα από την y_i .

Σχέση linear discriminant function και logit:

Η linear discriminant function αποτελεί στην ουσία μια ειδική περίπτωση του logit. Αν καλύπτονται οι υποθέσεις της discriminant function (αν οι παρατηρήσεις για τις οποίες το $y_i = 1$ και $y_i = 0$ είναι multivariate normal populations με μέσους M_1 και M_0 και covariance matrices $S_1 = S_0 = S$), τότε έχουν την ίδια μορφή συνάρτησης. Το αντίθετο όμως δεν ισχύει, καθώς το logit δεν υποθέτει πολυμεταβλητή κανονική κατανομή για τις ανεξάρτητες μεταβλητές.

Δείγμα: ως δείγμα χρησιμοποιήσε 58 αποτυχημένες τράπεζες.

Χρονική περίοδος: η χρονική περίοδος που τον ενδιέφερε ήταν το 1970-1976 και οι λογιστικές καταστάσεις αφορούσαν 2 χρόνια πριν την αποτυχία

Μεταβλητές: χρησιμοποίησε 25 δείκτες που κατηγοριοποιούνται ως δείκτες ρευστότητας, κεφαλαιακής επάρκειας, κερδοφορίας και asset risk. Κατόπιν ποικίλων συνδυασμών κατέληξε σε ένα μοντέλο με 4 δείκτες των τριών τελευταίων κατηγοριών. Οι δείκτες αφορούσαν το έτος 1974 για πρόβλεψη για τη χρονική περίοδο 1975-76. Οι εκτιμήσεις του logit ελέγχθηκαν έναντι της μηδενικής υπόθεσης ότι όλες οι τράπεζες είχαν την ίδια πιθανότητα αποτυχίας, η οποία εκτιμήθηκε σε 0,0041 (23 αποτυχημένες τράπεζες / 5598 σύνολο πληθυσμού). Αν πιθανότητα τράπεζας $> 0,0041 \rightarrow$ αποτυχημένη

Classification accuracy: Το logit παρουσίασε ακρίβεια κατάταξης 87% για τις αποτυχημένες τράπεζες και 88,6% για τις υγιείς.

Σύγκριση logit – discriminant analysis:

i) σε όρους classification accuracy: παρουσίασαν την ίδια ακρίβεια κατάταξης, με την discriminant analysis να είναι απλούστερη σε υπολογισμούς αλλά να θεωρείται ύποπτη λόγω της ανισότητας των covariance matrices και της μη κανονικότητας.

ii) σε όρους probability estimation: το logit παρουσιάστηκε καλύτερο.

Περαιτέρω έρευνα πρόβλεψης για τα έτη 1971-72 με δείκτες του 1970 έδειξε ότι η τετραγωνική δομή είναι καλύτερη και ότι η ακρίβεια κατάταξης και η πρόβλεψη των πιθανοτήτων μειώθηκαν τόσο για το ένα μοντέλο όσο και για το άλλο.

e) Newer models

Τα μονομεταβλητά και πολυμεταβλητά μοντέλα που αναλύσαμε έως τώρα έχουν υποστεί πολλές και ποικίλες κριτικές κυρίως επειδή τα περισσότερα από αυτά είναι γραμμικά και θεωρούνται ότι δεν ανταποκρίνονται στο μη γραμμικό κόσμο μας, επειδή βασίζονται σε λογιστικά στοιχεία και μπορεί έτσι να μην λάβουν υπόψη τις γρήγορα μεταβαλλόμενες συνθήκες του δανειζομένου και επειδή είναι πολύ λίγο συνδεδεμένα με ένα βασικό θεωρητικό μοντέλο. Δημιουργήθηκαν έτσι καινούργια μοντέλα πρόβλεψης πιστωτικού κινδύνου από τα οποία τα σημαντικότερα θεωρούνται τα εξής:

- i) risk of ruin models : τα μοντέλα αυτά θεωρούν ότι μια εταιρία χρεωκοπεί όταν η αξία του ενεργητικού της γίνεται μικρότερη από τις υποχρεώσεις της στους εξωτερικούς πιστωτές της. Τα μοντέλα αυτά μοιάζουν αρκετά στο Black-Scholes Merton model. KMV: το KMV είναι το σημαντικότερο από αυτά τα μοντέλα. Το θεωρητικό μοντέλο στο οποίο βασίζεται είναι το OPM model που θεωρεί την μετοχική αξία της επιχείρησης ως ένα call option στο ενεργητικό της. Οι αμφιβολίες όμως που υπάρχουν για το κατά πόσον η μεταβλητότητα της μετοχής μπορεί να δώσει την μεταβλητότητα του ενεργητικού της επιχείρησης και για το αν το KMV ισχύει για εταιρίες μη εισηγμένες στο χρηματιστήριο κάνουν το μοντέλο λιγότερο πρακτικό.
- ii) neural networks : μοιάζουν στην discriminant analysis αλλά δεν υιοθετούν την υπόθεση ότι οι μεταβλητές είναι γραμμικά και αυτόνομα εξαρτημένες. Τα μειονεκτήματα και πλεονεκτήματα αυτής της μεθόδου έχουν ήδη αναφερθεί νωρίτερα στην εργασία μας.
- iii) fixed income portfolio analysis : η μέθοδος αυτή προσπαθεί να αποφύγει τα προβλήματα των δεδομένων και της ανάλυσης. Αφορά ουσιαστικά ομολογίες αλλά μπορεί να εφαρμοστεί και σε δάνεια. Η γενική προσέγγιση στην οποία βασίζεται είναι η διαφοροποίηση του χαρτοφυλακίου.

3.8.ΙΔΙΑΙΤΕΡΟΤΗΤΕΣ ΜΟΝΤΕΛΩΝ ΠΙΣΤΩΤΙΚΟΥ ΚΙΝΔΥΝΟΥ

3.8.1.ΕΜΠΕΙΡΙΚΑ ΕΝΑΝΤΙ ΘΕΩΡΗΤΙΚΩΝ ΜΟΝΤΕΛΩΝ

Ο Scott το 1980, δημοσίευσε ένα άρθρο με θέμα τη σύγκριση μεταξύ θεωρητικών και εμπειρικών μοντέλων. Κατά το άρθρο αυτό, τα εμπειρικά μοντέλα υστερούν σε θεωρητική υποστήριξη και δίνουν έμφαση στον τύπο των μεταβλητών και στη λειτουργική μορφή των μοντέλων. Τα εμπειρικά μοντέλα διακρίνονται σε μοντέλα χωρίς θεωρητική υποστήριξη και μοντέλα με θεωρητική υποστήριξη.

A) Εμπειρικά μοντέλα

- i) *Μοντέλα χωρίς θεωρητική υποστήριξη*: τα μοντέλα αυτά βασίζονται μόνο σε εμπειρική δουλειά η οποία ακολουθεί κάποια συγκεκριμένα

βήματα: την επιλογή ενός αντιστοιχισμένου δείγματος, την επιλογή δεικτών, τη δημιουργία μιας συνάρτησης από δείκτες που διαχωρίζουν, την εφαρμογή στατιστικών ελέγχων της συνάρτησης και τον συνεχή έλεγχο και αναθεώρηση του μοντέλου μέσα στο χρόνο. Οι πιο γνωστές από τις εμπειρικές αυτές μελέτες έχουν ήδη αναφερθεί, με πρωτοπόρες του Beaver και του Altman. Όλα τα παραπάνω μοντέλα περιείχαν δείκτες μετοχών και ταμειακών ροών.

- ii) *Μοντέλα με θεωρητική υποστήριξη*: το πιο γνωστό εμπειρικό μοντέλο με θεωρητική υποστήριξη είναι το gambler's ruin theory που εφαρμόστηκε όπως προαναφέραμε χωρίς ιδιαίτερη επιτυχία από τους Wilcox, Santomero και Vinsco. Το μοντέλο αυτό βασίζεται στη θεωρία ότι η εταιρία χρεωκοπεί όταν η καθαρή της αξία γίνει μηδέν. Η θεωρία αυτή όμως αποδείχτηκε πολύ απλή καθώς υποθέτει ότι οι ταμειακές ροές προέρχονται από ανεξάρτητες δοκιμές χωρίς να λαμβάνει υπόψη την επέμβαση της διοίκησης.

Συμπεράσματα για τα εμπειρικά μοντέλα: η πρόβλεψη χρεωκοπίας με αυτά τα μοντέλα είναι δυνατή. Παρόλο που δεν έχουν θεωρητικό υπόβαθρο, τα συγκεκριμένα μοντέλα έχουν αποδειχτεί επιτυχή. Τα μοντέλα αυτά περιλαμβάνουν μεταβλητές κερδών, ταμειακών ροών και χρέους που βασίζονται σε λογιστικά στοιχεία και στοιχεία μετοχών. Τα πολυμεταβλητά εμπειρικά μοντέλα αποδείχτηκαν καλύτερα από τα μονομεταβλητά με το ZETA μάλλον να υπερέχει. Παρ' ολ' αυτά, έκπληξη προκάλεσε η πολύ καλή απόδοση των μονομεταβλητών μοντέλων που αποδόθηκε στην έλλειψη του προβλήματος overfitting(η επιλογή ενός set 4-5 μεταβλητών ανάμεσα σε 25, αριθμός που μπορεί να αρμόζει στο δείγμα αλλά όχι στον πληθυσμό) που παρουσιάζουν εν αντιθέσει τα πολυμεταβλητά. Ο κίνδυνος των πολυμεταβλητών μοντέλων στην περίπτωση αυτή έγκειται στην μη ακριβή πρόβλεψη με βάση νέα στοιχεία, κάτι που όμως αντιμετωπίζεται με μια σειρά από tests σε διαφορετικά δείγματα.

B)Θεωρητικά μοντέλα

- i) *Single-period models*: θεωρούν ότι η ζωή της επιχείρησης διαρκεί 2 περιόδους και ότι η χρεωκοπία επέρχεται όταν η ρευστοποίησημη αξία της γίνεται μικρότερη από τα χρέη. Δεν έχει καμία σχέση με τα εμπειρικά μοντέλα και δεν στηρίζεται σε λογιστικά στοιχεία και στοιχεία κερδών αλλά μόνο σε δεδομένα μετοχών.
- ii) *Gambler's ruin models*: θεωρούν ότι η εταιρία διαθέτει ένα δεδομένο ποσό κεφαλαίου K και χρεωκοπεί όταν $K < 0$. Το K αντιπροσωπεύει ουσιαστικά τη ρευστοποιήσιμη αξία του ενεργητικού. Η εταιρία θεωρείται ότι μπορεί να καλύψει τις ζημιές τις μόνο πουλώντας στοιχεία του ενεργητικού της, χωρίς να μπορεί να προσφύγει στην πώληση χρεογράφων. Το μοντέλο εξηγεί τα εμπειρικά μοντέλα αν θέσουμε το K ως ισοδύναμο με την λογιστική αξία της equity της εταιρίας και μοιάζει μάλιστα αρκετά με το ZETA.
- iii) *Perfect access to external capital*: η μετοχή εδώ θεωρείται ως πηγή κεφαλαίου στην οποία μπορεί να στηριχτεί η εταιρία σε περίπτωση

χρεωκοπίας όπου η market value of equity < 0 . Είναι μια δυναμική θεωρία που λαμβάνει υπόψη τις μεταβλητές εσωτερικής κερδοφορίας και των μετοχών. Τα εμπειρικά αποτελέσματα υποστηρίζουν το συγκεκριμένο μοντέλο.

Το γενικό συμπέρασμα του άρθρου είναι η διαπίστωση της ανάγκης για έναν συνδυασμό θεωρητικών και εμπειρικών μοντέλων προς ένα καλύτερο και επιστημονικά τεκμηριωμένο τελικό αποτέλεσμα. Θεωρήσαμε σκόπιμο να ανφερθούμε σε αυτόν τον ανταγωνισμό που υπάρχει ανάμεσα στα θεωρητικά και τα εμπειρικά μοντέλα και στην τάση που εκδηλώνεται για μια αποδοτικότερη προσέγγιση του προβλήματος μέσω του συνδυασμού τους.

3.8.2. BIAS

Ένα ακόμη θέμα που αξίζει να αναφερθεί είναι οι παραβιάσεις που εμφανίζονται στα μοντέλα πρόβλεψης πιστωτικού κινδύνου εξαιτίας του δείγματος και των δεδομένων. Μία έρευνα που έγινε το 1984 από τον Zmijewski αναφέρεται σε δύο ειδών biases:

- *sample selection bias* : Η πρώτη, που αφορά το δείγμα, προκαλείται από το γεγονός ότι τα μοντέλα αυτά εφαρμόζονται σε μη τυχαία δείγματα. Το δείγμα δηλαδή που επιλέγεται περιέχει πολύ περισσότερες χρεωκοπημένες εταιρίες από ότι θα έπρεπε αν λάβουμε υποψη τον πληθυσμό. Τα περισσότερα δείγματα που επιλέχθηκαν περιείχαν μια αντιστοίχιση 1 προς 1 υγιών με προβληματικές εταιρίες, π.χ. 50 υγιείς με 50 προβληματικές. Σε έναν κανονικό πληθυσμό όμως, η αντιστοίχιση αυτή δεν ισχύει. Υπάρχουν πολύ λιγότερες προβληματικές εταιρίες από ότι υγιείς με λόγο π.χ. 29 προς 5500 αντίστοιχα. Ενώ λοιπόν στον πληθυσμό υπάρχει μια τέτοια αναλογία προβληματικών-υγιών, στο δείγμα λαμβάνεται μια αναλογία με τις προβληματικές να καταλαμβάνουν το 50% του δείγματος κάτι που δεν ανταποκρίνεται στην πραγματικότητα. Οι περισσότερες εκτιμητικές τεχνικές υποθέτουν τη χρήση ενός τυχαίου εξωγενούς δείγματος στο οποίο κάθε παρατήρηση λαμβάνεται τυχαία ενώ οι εξαρτημένες και ανεξάρτητες μεταβλητές παρατηρούνται. Σε αυτά τα μοντέλα αντίθετα, που είναι *choise-based sample*, η πιθανότητα να περιληφθεί μια παρατήρηση στο δείγμα εξαρτάται από την αξία της εξαρτημένης μεταβλητής. Πρώτα δηλαδή ο αναλυτής παρατηρεί την εξαρτημένη μεταβλητή (π.χ. χρεωκοπημένη εταιρία) και μετά επιλέγει το δείγμα, βασίζοντάς το στο παραπάνω δεδομένο. Κάτι τέτοιο όμως παραβιάζει την υπόθεση της τυχαίας δειγματοληψίας και προκαλεί ασυμπτωτικό bias των εκτιμήσεων των πιθανοτήτων και παραμέτρων. Όταν ένα group εξαρτημένων μεταβλητών έχει δείγμα με πιθανότητα μεγαλύτερη από αυτή του πληθυσμού, είναι *oversampled* και παρουσιάζει υποεκτίμηση των *classification* και *prediction error rates*. Η μελέτη του Zmijewski έδειξε ότι η πραγματική συχνότητα χρεωκοπίας στον πληθυσμό είναι μικρότερη του 0,75%. Η συχνότητα όμως που προβλέπεται από τα μοντέλα κυμαίνεται μεταξύ του 1,5% και του 50%! Βέβαια τα *choise-based sample designs* δεν είναι κατ'ανάγκη ακατάλληλα. Δείγματα που συλλέγονται με αυτόν τον τρόπο είναι συνήθως μη

δαπανηρά σε αντίθεση με τα τυχαία δείγματα που περιέχουν τη δύσκολη και δαπανηρή διαδικασία της ένταξης ενός ικανού αριθμού σπάνιων γεγονότων και παρατηρήσεων για την εκτίμηση του μοντέλου.

- *data collection bias*: Το δεύτερο bias προέρχεται από την έλλειψη επαρκών και ακριβών στοιχείων για προβληματικές εταιρίες. Η ανεπάρκεια αυτή οφείλεται στην πολύ μικρή ηλικία των εταιριών που αποτυγχάνουν (πράγματι οι περισσότερες προβληματικές εταιρίες είναι νέες εταιρίες) και επιπλέον στην άρνηση τους να δημοσιοποιήσουν αυτά τα στοιχεία. Αν εκτιμηθεί ένα μοντέλο με τη δέσμευση της επάρκειας στοιχείων και αγνοηθεί η πραγματική έλλειψη δεδομένων, οι πιθανότητες που θα υπολογιστούν θα υποεκτιμούν τις πιθανότητες του πληθυσμού. Θα υπάρχουν δηλαδή πολλές προβληματικές εταιρίες με ελλιπή στοιχεία στον πληθυσμό που δεν θα προβλεφθούν αφού δεν θα έχουν ληφθεί υπόψη αρχικά από το μοντέλο.

Οι 2 παραπάνω *sample selection* και *data collection* περιορισμοί λοιπόν που τίθενται αναπόφευκτα στα μοντέλα δημιουργούν biases που πρέπει να ληφθούν υπόψη. Όσον αφορά το πρόβλημα που προκύπτει από την μη τυχειότητα του δείγματος, αυτό αντιμετωπίζεται με κάποιες εναλλακτικές και προσαρμοσμένες τεχνικές για την εκτίμηση των παραμέτρων (*weighted exogenous sample maximum likelihood, conditional maximum likelihood, full information concentrated maximum likelihood*). Το πρόβλημα από την άλλη που προκύπτει από την έλλειψη στοιχείων προβληματικών εταιριών μπορεί να αντιμετωπιστεί αν υποθεθεί ότι οι εταιρίες με ανεπαρκή στοιχεία αποτελούν τυχαίο δείγμα.

Παρολαυτά, οφείλουμε να επισημάνουμε ότι το βασικό πρόβλημα που προκύπτει από τα 2 αυτά biases αφορά τη συχνότητα εμφάνισης προβληματικών εταιριών μέσα στο σύνολο του πληθυσμού κάτι που την προκείμενη τουλάχιστον διπλωματική εργασία αφήνει ανεπηρέαστη. Από τη στιγμή δηλαδή που τα συνολικά *prediction* και *classification rates* δεν επηρεάζονται, τα συγκεκριμένα biases δεν χρειάζεται να ληφθούν υπόψη. Εμάς μας ενδιαφέρει κατά πόσον το κάθε μοντέλο είναι ικανό να προβλέψει την εμφάνιση προβλήματος σε μια μεμονωμένη εταιρία και όχι τη συχνότητα της εμφάνισης προβλήματος σε όλον το πληθυσμό. Άλλωστε τα biases αποδείχτηκαν να έχουν τόσο μικρή επιρροή στην ακρίβεια των μοντέλων που δεν θα έπρεπε να μας απασχολούν. Παρολαυτά θεωρήθηκε σκόπιμο να αναφερθούμε και σε αυτήν την πτυχή του θέματος.

1. ΣΤΟΧΟΣ ΕΡΓΑΣΙΑΣ

Σκοπός αυτής της εργασίας είναι όπως έχουμε προαναφέρει η εφαρμογή και σύγκριση ορισμένων οικονομετρικών μοντέλων πιστωτικού κινδύνου σε δεδομένα της ελληνικής αγοράς. Από τα τέσσερα μοντέλα που αναλύσαμε στο θεωρητικό μέρος αυτής της εργασίας, καταλήξαμε στην τελική χρήση και σύγκριση 3 τέτοιων μοντέλων, του logit του probit και της MDA. Το Linear Probability Model απορρίφθηκε τελικά ως μία γενική μορφή παλινδρόμησης, με αρκετά μειονεκτήματα όπως αυτά που αναφέρθηκαν στο θεωρητικό μέρος και που αφορούσαν την ετεροσκεδαστικότητα και την επέκταση των αποτελεσμάτων εκτός των ορίων (0,1). Θεωρούμε ότι τα αποτελέσματα που θα μας έδινε ένα τέτοιο μοντέλο θα ήταν αν όχι λιγότερο ακριβή, τουλάχιστον παρόμοια με αυτά των υπολοίπων.

Η διαδικασία που ακολουθήσαμε μέχρι τα τελικά αποτελέσματα της έρευνας, διακρίνεται σε πολλά επιμέρους στάδια:

2. ΟΡΙΣΜΟΣ ΠΡΟΒΛΗΜΑΤΟΣ

Η σκοπιμότητα των μοντέλων πιστωτικού κινδύνου έγκειται στη δυνατότητα πρόβλεψης της χρεωκοπίας μιας επιχείρησης. Αν για παράδειγμα μια τράπεζα ενδιαφέρεται να δανειοδοτήσει μια εταιρία θα χρησιμοποιήσει ένα τέτοιο μοντέλο ως ένα επιμέρους κριτήριο, για να προβλέψει κατά πόσον η εταιρία αυτή θα εξακολουθήσει να είναι υγιής ή μη στο άμεσο μέλλον και επομένως πόσο πιθανό είναι να αθετήσει τις υποχρεώσεις της προς την τράπεζα ή όχι. Οι προηγούμενες μελέτες που έχουν διεξαχθεί πάνω στο συγκεκριμένο θέμα αφορούσαν κυρίως τη χρεωκοπία της επιχείρησης. Στη δική μας διατριβή η προβλεψιμότητα των μοντέλων θα περιορίζεται στην καθυστέρηση ή μη των επιχειρήσεων. Ο ελληνικός νόμος διαφέρει από τους ξένους όσον αφορά το θέμα της πτώχευσης των εταιριών και κηρύττει σε πτώχευση επιχειρήσεις που βρίσκονται στο τελευταίο στάδιο της επιβίωσης τους παρόλο που έχουν προ πολλού αρχίσει να αθετούν τις υποχρεώσεις τους. Μια τράπεζα όμως θα την ενδιέφερε πάνω απ' όλα η καθυστέρηση και η συχνά επακόλουθη αθέτηση των δανειακών υποχρεώσεων της επιχείρησης παρά η καθεαυτή χρεωκοπία της. Επικεντρώνουμε επομένως το πρόβλημά μας στην **αθέτηση των υποχρεώσεων** μιας επιχείρησης και όχι στη χρεωκοπία της. Γι' αυτόν ακριβώς το λόγο χρησιμοποιούμε δεδομένα υγιών και καθυστερημένων επιχειρήσεων για τη δημιουργία και εφαρμογή των μοντέλων μας.

Κατ' αυτόν τον τρόπο και η διττή εξαρτημένη μεταβλητή που περιέχει κάθε μοντέλο ορίζεται από τις τιμές 0 και 1 όπου η 0 αντιστοιχεί σε καθυστερημένη επιχείρηση και η 1 σε υγιή.

3. ΑΝΤΙΣΤΟΙΧΙΣΗ-ΔΗΜΙΟΥΡΓΙΑ ΔΕΙΓΜΑΤΩΝ

Η επιλογή δείγματος αποτελεί το σημαντικότερο ίσως στάδιο μιας μελέτης, καθώς τα στοιχεία που θα τροφοδοτήσουν τα οικονομετρικά προγράμματα θα καθορίσουν και την ποιότητα των εξαγόμενων αποτελεσμάτων.

Προβλήματα: Σε προηγούμενες μελέτες που έχουν γίνει κυρίως στο εξωτερικό, το κάθε δείγμα προκύπτει από την αντιστοίχιση υγιών -προβληματικών εταιριών με κριτήρια τον κλάδο, το μέγεθος και τα έτη. Η περιορισμένου μεγέθους ελληνική αγορά όμως αδυνατεί να προσφέρει επάρκεια σε τέτοιου είδους στοιχεία. Δεν θα μπορούσαμε για παράδειγμα να αντιστοιχίσουμε 30 υγιείς με 30 καθυστερημένες τεχνικές εταιρίες γιατί απλούστατα στην χώρα μας δεν υφίσταται τόσο μεγάλος αριθμός καθυστερημένων εταιριών μέσα σε έναν κλάδο. Αλλά ακόμη και αν υποθέσουμε ότι υφίσταται θα βρισκόμασταν πάλι αντιμέτωποι με την άρνηση αυτών των εταιριών να δημοσιοποιήσουν τα στοιχεία τους. Το βασικό

δυστυχώς πρόβλημα που αντιμετωπίσαμε στη δική μας μελέτη ήταν αυτή ακριβώς η ανεπάρκεια ολοκληρωμένων οικονομικών δεδομένων. Με βάση τα στοιχεία που διαθέταμε και που αφορούσαν εταιρίες μικρομεσαίου μεγέθους, βιοτεχνικού και εμπορικού κλάδου και διαφόρων ετών από το 1996 μέχρι και το 2001, προσπαθήσαμε να επιτύχουμε την καλύτερη δυνατή αντιστοίχιση.

Βάση δεδομένων: Τα αρχικά μας στοιχεία, που προήλθαν από τη βάση δεδομένων μεγάλης ελληνικής τράπεζας, αφορούσαν 340 επιχειρήσεις, 204 υγιείς και 136 καθυστερημένες. Οι τελευταίες δηλώθηκαν σε καθυστέρηση κατά τη διάρκεια των ετών 1997-2001. Λαμβάνοντας υπόψη την ανομοιογένεια του γκρουπ και τις διαφορές κλάδου και μεγέθους υπήρξαμε πολύ προσεκτικοί στην αντιστοίχιση. Από όλες τις παραπάνω μπορέσαμε και καταλήξαμε σε ένα αρχικό δείγμα 84 εταιριών, 42 υγιών-42 καθυστερημένων, κρίνοντας κάθε φορά την ομοιότητά τους σύμφωνα με τον κλάδο, το μέγεθός τους, δηλαδή τον κύκλο εργασιών τους, και τα έτη για τα οποία διαθέταμε στοιχεία..

Αντιστοίχιση βάσει κλάδου και αντικειμένου εργασιών: Αντιστοιχίσαμε με άλλα λόγια βιοτεχνικές υγιείς με βιοτεχνικές καθυστερημένες επιχειρήσεις και εμπορικές υγιείς με εμπορικές καθυστερημένες.. Η ομοιότητα όμως των επιχειρήσεων δεν περιορίστηκε μόνο στον κλάδο αλλά και στο επιμέρους αντικείμενο παραγωγής ή πώλησης της κάθε εταιρίας. Φροντίσαμε δηλαδή μια υγιής βιοτεχνία παραγωγής ενδυμάτων για παράδειγμα να αντιστοιχηθεί με μία προβληματική βιοτεχνία που να παράγει επίσης ενδύματα.

Αντιστοίχιση βάσει μεγέθους: Το μέγεθος των εταιριών του δείγματος ποικίλλει ξεκινώντας από 200 εκ. μέχρι και 1 δις δρχ. κύκλο εργασιών. Κατά τον ίδιο τρόπο υγιής εταιρία του ίδιου πια κλάδου και αντικειμένου και ορισμένου κύκλου εργασιών αντιστοιχήθηκε σε καθυστερημένη εταιρία ομοίου κύκλου εργασιών.

Αντιστοίχιση βάσει ετών: Όσον αφορά τα έτη, για τις προβληματικές εταιρίες της αρχικής βάσης δεδομένων διαθέταμε στοιχεία από το 1996 έως και το 2000. Επειδή όμως για τις υγιείς τα στοιχεία μας περιορίζονταν στα έτη 1997-2000 αναγκαστικά η αντιστοίχισή μας κυμάνθηκε στα τελευταία αυτά τρία έτη αφήνοντας απ' έξω το 1996. Το δείγμα τελικά στο οποίο καταλήξαμε περιέχει στοιχεία επιχειρήσεων μέχρι και 2 ετών πριν την καθυστέρησή τους. Πιο συγκεκριμένα, για 2 έτη πριν την καθυστέρηση είχαμε στοιχεία για τις 54 (27-27) από τις 84 επιχειρήσεις, ενώ για 1 έτος πριν και για τις 84(42-42). Οι 54 εταιρίες για τις οποίες είχαμε δεδομένα για 2 έτη πριν, δεν καθυστέρησαν απαραίτητα την ίδια χρονιά. Τα 2 αυτά έτη προ καθυστέρησης ήταν ή το 1997 και το 1998 για τις εταιρίες που καθυστέρησαν το 1999 ή το 1998 και το 1999 για αυτές που καθυστέρησαν το 2000. Αντίστοιχα, τα δεδομένα των εταιριών 1 έτους πριν την καθυστέρηση αφορούν τη χρονιά 1998 για τις εταιρίες που καθυστέρησαν το 1999 και τη χρονιά 1999 για εκείνες που καθυστέρησαν το 2000.

Ξεκινώντας έτσι από τον κλάδο, συνεχίζοντας με το μέγεθος και συγκρίνοντας τέλος τα έτη, καταλήξαμε σε ένα αρχικό δείγμα 84 αντιστοιχισμένων, με τον καλύτερο δυνατό τρόπο, επιχειρήσεων. Από το αρχικό αυτό δείγμα δημιουργήσαμε τα μοντέλα μας χρησιμοποιώντας όμως μόνο τα στοιχεία των 84 εταιριών 1 έτους πριν την καθυστέρηση. Θεωρήσαμε ότι το μέγεθος αυτού του συγκεκριμένου δείγματος ήταν επαρκές για τη δημιουργία μοντέλων (42 ζεύγη παρατηρήσεων). Ο κυριότερος όμως λόγος για τη χρήση αυτού μόνο του δείγματος και όχι ολόκληρου, δηλαδή και αυτού των 2 ετών, έγκειται στο ότι θεωρούμε ότι το δείγμα που

βρίσκεται όσο το δυνατόν πιο κοντά χρονικά στην καθυστέρηση θα περιέχει πιο ξεκάθαρα και ενδεικτικά στοιχεία και θα μας δώσει πιο ακριβές μοντέλο πρόβλεψης. Θεωρούμε με άλλα λόγια ότι τα οικονομικά στοιχεία μιας προβληματικής εταιρίας 1 έτους πριν την καθυστέρηση διαφέρουν πολύ περισσότερο από αυτά των 2 ή 3 ετών πριν, σε σχέση με τα αντίστοιχα στοιχεία μιας υγιούς εταιρίας, αποτελώντας έτσι σοβαρές ενδείξεις της επερχόμενης κατάληξης της εταιρίας.

Το υπόλοιπο του αρχικού δείγματος (54 παρατηρήσεις για 2 έτη πριν) χρησιμοποιήθηκε για τη διεξαγωγή διαφόρων τεστ ελέγχου της ακρίβειας και της προβλεψιμότητας του μοντέλου. Για τον ίδιο σκοπό, χρησιμοποιήσαμε και ένα δεύτερο δείγμα, που δημιουργήθηκε για μια ακόμη φορά με τους κανόνες αντιστοίχισης. Το καινούργιο αυτό δείγμα όμως των 42 αντιστοιχισμένων εταιριών (21 υγιών-21 καθυστερημένων), προέκυψε από μία αντιστοίχιση βάσει μόνο μεγέθους και ετών. Δεν λάβαμε υπόψη μας δηλαδή καθόλου τον κλάδο, κι αυτό γιατί κάθε δυνατή αντιστοίχιση βάσει αυτού του κριτηρίου είχε εξαντληθεί στο αρχικό δείγμα. Την ανεπάρκεια αυτή του δείγματος οφείλουμε να τη λάβουμε υπόψη κατά τη μελέτη των αποτελεσμάτων των τεστ ελέγχου.

4.ΔΗΜΙΟΥΡΓΙΑ ΜΟΝΤΕΛΩΝ

Τα οικονομετρικά προγράμματα που χρησιμοποιήσαμε για τη δημιουργία και εφαρμογή των τριών μοντέλων ήταν το E-views για τα logit και probit και το SPSS για την MDA.

Γενική μορφή: Και τα τρία μοντέλα, παρά τις μικροδιαφορές που έχουν σε θεωρητικό επίπεδο και που έχουν ήδη αναφερθεί στο θεωρητικό μέρος αυτής της εργασίας, έχουν την εξής γενική μορφή:

$$y = c + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \dots + \beta_k x_k$$

Εξαρτημένη μεταβλητή: Ορίσαμε ήδη τη διττή μεταβλητή y να λαμβάνει τιμές 0 όταν αντιστοιχεί σε καθυστερημένη επιχείρηση και 1 όταν αντιστοιχεί σε υγιή.

Ανεξάρτητες μεταβλητές: Οι ανεξάρτητες μεταβλητές x αποτελούν χρηματοοικονομικούς δείκτες των εταιριών, οι οποίοι προέκυψαν από τα οικονομικά δεδομένα που διαθέταμε, με κατάλληλες αριθμητικές πράξεις. Με βάση προηγούμενες ελληνικές και ξένες μελέτες και τη δική μας κρίση επιλέξαμε 13 χρηματοοικονομικούς δείκτες “υποψήφιους” για την ένταξή τους στα μοντέλα μας, των κατηγοριών ρευστότητας, δραστηριότητας, αποδοτικότητας και βιωσιμότητας:

x_1 : Κυκλοφορούν Ενεργητικό / Βραχυπρόθεσμες Υποχρεώσεις : Δείχνει τη ρευστότητα των Βραχυχρόνιων Υποχρεώσεων της επιχείρησης και την ικανότητα των κυκλοφοριακών στοιχείων της να καλύψουν τις παραπάνω υποχρεώσεις. Ένας σχετικά υψηλός αριθμοδείκτης παρέχει ένδειξη ότι, αν δεν υπάρξουν υπερβολικές ζημιές από τη ρευστοποίηση των κυκλοφοριακών στοιχείων, οι Βραχυπρόθεσμες Υποχρεώσεις θα μπορέσουν να εξοφληθούν από τα κεφάλαια κινήσεως. Αντίθετα, ένας μικρός αριθμοδείκτης δείχνει επιδείνωση της κατάστασης της εταιρίας από λειτουργικές και έκτακτες ζημιές.

x_2 : Κυκλοφορούν Ενεργητικό / Ενεργητικό : Αποτελεί ένα μέτρο των στοιχείων του ενεργητικού της επιχείρησης που μπορούν να ρευστοποιηθούν σε σχέση με τη συνολική της κεφαλαιοποίηση. Μία προβληματική εταιρία που βιώνει λειτουργικές

ζημιές θα παρουσιάζει συρρικνωμένο και διαρκώς μειούμενο Κυκλοφορούν Ενεργητικό σε σχέση με το Συνολικό Ενεργητικό της.

x₃: Παθητικό / Ενεργητικό : Δείχνει την ικανότητα της επιχείρησης να ικανοποιεί το σύνολο των υποχρεώσεων της από τα στοιχεία του Ενεργητικού της. Σε προβληματική κατάσταση τον αναμένουμε αυξημένο.

x₄: Παθητικό / Καθαρή Θέση : Δείχνει πόσο μπορούν τα περιουσιακά στοιχεία της επιχείρησης που μετριούνται με την αξία της Καθαρής Θέσης συν το χρέος, να μειωθούν σε αξία προτού οι υποχρεώσεις της υπερβούν το ενεργητικό της και η εταιρία κριθεί μη υγιής. Σε μία προβληματική εταιρία θα περιμέναμε να δούμε αυτόν τον δείκτη αυξημένο, καθώς η Καθαρή της θέση δεν θα μπορούσε πλέον να καλύψει το Συνολικό Παθητικό της.

x₅: Βραχυπρόθεσμες Υποχρεώσεις / Παθητικό : Αποτελεί ένα μέτρο της ρευστότητας των υποχρεώσεων της επιχείρησης, μετρώντας την αναλογία του Συνολικού Παθητικού που αντιπροσωπεύουν οι Βραχυπρόθεσμες Υποχρεώσεις. Όσο πιο μεγάλος είναι ο δείκτης, δηλαδή όσο πιο μεγάλο μέρος του συνολικού Παθητικού αποτελούν οι Βραχυπρόθεσμες Υποχρεώσεις, τόσο πιο δεινή είναι η κατάσταση της επιχείρησης, αφού οφείλει πολύ σύντομα να εξοφλήσει μεγάλο αριθμό υποχρεώσεων.

x₆: Κεφάλαιο Κίνησης / Ενεργητικό : Το Κεφάλαιο Κίνησης ορίζεται ως η διαφορά μεταξύ Κυκλοφορούντος Ενεργητικού και Βραχυπρόθεσμων Υποχρεώσεων. Ο δείκτης αυτός επομένως αποτελεί ένα μέτρο καθαρής ρευστότητας της επιχείρησης σε σχέση με το Συνολικό Ενεργητικό της. Ένας τέτοιος δείκτης προβληματικής εταιρίας θα αναμέναμε να μειώνεται.

x₇: Κέρδη προ φόρων / Ενεργητικό : Αποτελεί έναν δείκτη κερδοφορίας που δείχνει την ικανότητα της επιχείρησης να παράγει κέρδη από τη χρήση του Ενεργητικού της. Και αυτόν σε μία μη υγιή επιχείρηση θα τον περιμέναμε μειωμένο.

x₈: Κέρδη προ φόρων / Καθαρή Θέση : Πρόκειται για έναν δείκτη αποδοτικότητας των κεφαλαίων της εταιρίας. Όσο πιο μεγάλος, τόσο καλύτερα για αυτήν.

x₉: Κέρδη προ φόρων / Βραχυπρόθεσμες Υποχρεώσεις : Αποτελεί μια ένδειξη του κατά πόσον τα κέρδη προ φόρων της επιχείρησης μπορούν να καλύψουν τις Βραχυπρόθεσμες Υποχρεώσεις της. Όσο πιο χαμηλός είναι, τόσο πιο προβληματική η κατάσταση της επιχείρησης.

x₁₀: Κέρδη μετά φόρων / Βραχυπρόθεσμες Υποχρεώσεις : Είναι ένα μέτρο της κερδοφορίας της επιχείρησης σε σχέση με τις Βραχυπρόθεσμες Υποχρεώσεις της. Ένας μειούμενος δείκτης θα αποτελούσε ένδειξη για τη δυσκολία της επιχείρησης να ικανοποιήσει τις Βραχυπρόθεσμες Υποχρεώσεις της με εσωτερικές πηγές χρηματοδότησης. Ίσως αποτελεί καλύτερο μέτρο από τον προηγούμενο δείκτη, καθώς είναι πιθανό τα κέρδη μετά τον υπολογισμό των φόρων να μην επαρκούν τελικά για την κάλυψη των υποχρεώσεων.

x₁₁: Πωλήσεις / Ενεργητικό : Αποτελεί ένα ενδεικτικό μέτρο της ικανότητας της εταιρίας να δημιουργεί Πωλήσεις από το Ενεργητικό της. Δείχνει επίσης την ικανότητα της διοίκησης να ανταποκρίνεται στις ανταγωνιστικές απαιτήσεις της αγοράς.

x_{12} : Δείκτης Αποδοτικότητας Ενεργητικού = Κέρδη μετά φόρων / Ενεργητικό : Δείχνει την ικανότητα του Ενεργητικού να παράγει Κέρδη. Όσο μεγαλύτερος είναι τόσο καλύτερα για την επιχείρηση.

x_{13} : Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων = Κέρδη μετά φόρων / Ίδια Κεφάλαια : Δείχνει την ικανότητα των Ιδίων Κεφαλαίων που διαθέτει η επιχείρηση να παράγουν Κέρδη. Και εδώ, ένας αυξημένος δείκτης αποτελεί μια πολύ θετική ένδειξη.

Διαδικασία κατασκευής: Στόχος αυτού του σταδίου της εμπειρικής εργασίας μας είναι η κατασκευή των μοντέλων. Για την επίτευξη αυτού του στόχου ακολουθήσαμε την εξής διαδικασία: εισάγαμε τιμές για τους 13 αριθμοδείκτες των 84 εταιριών καθώς και τιμές για τα αντίστοιχα y (42 μηδενικά και 42 μονάδες) και λύσαμε τις 84 αυτές εξισώσεις που δημιουργήθηκαν, ως προς τις παραμέτρους. Δώσαμε δηλαδή ως δεδομένα τα y και τα x και ζητήσαμε τα $c, \beta_1, \beta_2 \dots \beta_{13}$.

Τελική επιλογή ανεξάρτητων μεταβλητών:

Probit-logit: Επειδή όμως μας ενδιέφερε να καταλήξουμε στα μοντέλα με τις 5 καλύτερες ανεξάρτητες μεταβλητές, εφαρμόσαμε ένα iterative πρόγραμμα το οποίο με δεδομένα όλα τα παραπάνω μας έδωσε 1288 συνδυασμούς των 5 μεταβλητών για κάθε μοντέλο. Από αυτούς, και με κριτήρια τη στατιστική σημαντικότητα των μεταβλητών και τη συνολική ακρίβεια του κάθε συνδυασμού-μοντέλου, επιλέξαμε το βέλτιστο. Οι εναλλακτικές ανάμεσα στις οποίες έπρεπε να επιλέξουμε παρατίθενται στο παράρτημα στον ΠΙΝΑΚΑ 1 με μία αναλυτική παρουσίαση όλων των στατιστικών κριτηρίων που χρησιμοποιήσαμε για την τελική επιλογή.

Οι μεταβλητές στις οποίες καταλήξαμε είναι οι εξής:

Probit :

x_2 : Κυκλοφορούν Ενεργητικό / Ενεργητικό

x_6 : Κεφάλαιο Κίνησης / Ενεργητικό

x_9 : Κέρδη προ φόρων / Βραχυπρόθεσμες Υποχρεώσεις

x_{11} : Πωλήσεις / Ενεργητικό

x_{13} : Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων

Logit :

x_2 : Κυκλοφορούν Ενεργητικό / Ενεργητικό

x_6 : Κεφάλαιο Κίνησης / Ενεργητικό

x_{10} : Κέρδη μετά φόρων / Βραχυπρόθεσμες Υποχρεώσεις

x_{11} : Πωλήσεις / Ενεργητικό

x_{13} : Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων

Παρατηρούμε ότι τα δύο αυτά μοντέλα επιλέγονται ως αποτελεσματικότερες για πρόβλεψη σχεδόν τις ίδιες μεταβλητές με μια μικρή διαφορά στην επιλογή ανάμεσα στις x_9 και x_{10} . Ακόμη όμως και οι δύο αυτές μεταβλητές είναι περίπου οι ίδιες αφού πρόκειται για τα κέρδη προ και μετά φόρων. Οφείλουμε επίσης να πούμε ότι τα κριτήρια για την επιλογή της x_{10} στο μοντέλο logit ήταν πραγματικά ελάχιστα καλύτερα από της x_9 . Το αποτέλεσμα αυτό της ομοιότητας των αποτελεσμάτων των δύο αυτών μοντέλων ήταν αναμενόμενο από τη μελέτη και μόνο της θεωρητικής προσέγγισης που έγινε νωρίτερα στην εργασία μας, όπου αναφέραμε ότι θεωρητικά τα μοντέλα αυτά διαφέρουν ελάχιστα μεταξύ τους. Άλλωστε και η εμπειρία του

παρελθόντος έχει αποδείξει ότι σε μελέτες άλλων ερευνητών, τα δύο αυτά μοντέλα, κατέληγαν πάντα σε πολύ όμοια αποτελέσματα.

Αυτό που αξίζει ίσως να παρατηρήσουμε είναι ότι ως χρήσιμοι για την προβλεψιμότητα των μοντέλων πιστωτικού κινδύνου εμφανίζονται οι δείκτες κερδοφορίας, ρευστότητας και αποδοτικότητας με επικρατέστερους τους x_2 , x_6 και αμέσως μετά x_{11} και x_{13} , δηλαδή αντίστοιχα :

Κυκλοφορούν Ενεργητικό / Ενεργητικό, Κεφάλαιο Κίνησης / Ενεργητικό, Πωλήσεις / Ενεργητικό και Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων.

Οι μεταβλητές στις οποίες καταλήξαμε για την MDA είναι οι εξής:

MDA :

x_3 : Παθητικό / Ενεργητικό

x_{11} : Πωλήσεις / Ενεργητικό

x_{13} : Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων

Όπως βλέπουμε η MDA περιορίζει τους δείκτες σε 3, με τους δύο από αυτούς να συμπίπτουν με των δύο παραπάνω μοντέλων. Και πάλι έχουμε να κάνουμε με δείκτες ρευστότητας και αποδοτικότητας.

Οι εξισώσεις που υπολογίστηκαν για τα μοντέλα είναι οι παρακάτω:

$$\text{Probit : } y = - 2.576 * X_2 + 4.425 * X_6 + 0.262 * X_9 + 1.429 * X_{11} + 1.418 * X_{13}$$

$$\text{Logit : } y = - 4.382 * X_2 + 7.463 * X_6 + 0.432 * X_{10} + 2.521 * X_{11} + 2.383 * X_{13}$$

$$\text{MDA : } y = - 1.081 * X_3 + 0.842 * X_{11} + 0.649 * X_{13}$$

5.ΕΜΠΕΙΡΙΚΑ ΑΠΟΤΕΛΕΣΜΑΤΑ

5.1.Probit

5.1.1.Στατιστική σημαντικότητα μεταβλητών:

Πίνακας στατιστικών μέτρων μεταβλητών

Dependent Variable: Y				
Method: ML - Binary Probit				
Date: 05/11/02 Time: 15:32				
Sample(adjusted): 1 84				
Included observations: 84 after adjusting endpoints				
Convergence achieved after 4 iterations				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.174135	0.353124	-0.493127	0.6219
X ₂	-2.57617	0.855508	-3.011276	0.0026
X ₆	4.425213	1.052345	4.205097	0
X ₉	0.262444	0.116971	2.24366	0.0249
X ₁₁	1.42859	0.392584	3.638941	0.0003
X ₁₃	1.417527	0.483673	2.930753	0.0034

Ο παραπάνω πίνακας είναι το αποτέλεσμα που εξάγει το E-views όταν δίνουμε την εντολή να δημιουργήσει ένα μοντέλο probit με δεδομένα το y και τους παραπάνω 5 δείκτες.

Η πρώτη γραμμή του πίνακα μας πληροφορεί ποια είναι η εξαρτημένη μας μεταβλητή.

Η δεύτερη αναφέρεται στη μέθοδο που χρησιμοποιήσαμε για να επιλύσουμε την εξίσωση, που είναι η μέθοδος Maximum Likelihood μέσα στα πλαίσια του μοντέλου διττής εξαρτημένης μεταβλητής probit. Αν ανατρέξουμε με άλλα λόγια σε κάποιες αναφορές που έχουν γίνει στο θεωρητικό μέρος αυτής της εργασίας, θα θυμηθούμε ότι το μοντέλο probit όσο και το logit, συνεπάγονται την δημιουργία μιας Likelihood function της μορφής $L = \prod F(-\beta'x_i) \prod [1 - F(-\beta'x_i)]$, και μέσα στην οποία βρίσκουμε και τη διαφορά αυτών των δύο μοντέλων που προκύπτει από τις υποθέσεις που έχουμε κάνει για τη λειτουργική μορφή της αθροιστικής κατανομής των καταλοίπων F (αν είναι κανονική → probit, αν είναι λογαριθμική → logit). Η διαδικασία επίλυσης της εξίσωσης έγκειται στην εύρεση ενός σετ παραμέτρων που μεγιστοποιούν την L. Αυτή ακριβώς η διαδικασία ονομάζεται Maximum Likelihood estimation.

Η τρίτη γραμμή του πίνακα αναφέρεται στην ημερομηνία και την ώρα που εκτελέστηκε το πρόγραμμα.

Η τέταρτη γραμμή του πίνακα αναφέρεται στο μέγεθος του δείγματος, το οποίο αποτελείται όπως προαναφέραμε από δείκτες 84 εταιριών 1 έτους πριν την καθυστέρηση των προβληματικών.

Η πέμπτη γραμμή μας δείχνει ότι το πρόγραμμα συμπεριέλαβε όλες τις παρατηρήσεις στη διαδικασία, και τις 84.

Η έκτη γραμμή μας ενημερώνει για τον αριθμό των δοκιμών που έγιναν μέχρι να καταλήξει στις καλύτερους παραμέτρους. Στην περίπτωση αυτή το πρόγραμμα επίλυσε την εξίσωση με τον καλύτερο δυνατό τρόπο μετά από 4 δοκιμές. Οφείλουμε όμως να υπενθυμίσουμε ότι προηγουμένως και πριν καταλήξουμε στις 5 αυτές μεταβλητές, έχουμε ήδη “τρέξει” ένα άλλο iterative πρόγραμμα στο E-views το οποίο έκανε 1288 δοκιμές για όλες τις μεταβλητές. Επομένως το σύνολο των iterations που έχουν γίνει είναι στην ουσία πολύ μεγαλύτερο. Οι 4 δοκιμές αφορούν μόνο αυτές τις 5 μεταβλητές.

Variable: Η πρώτη στήλη του πίνακα περιέχει τις μεταβλητές του μοντέλου, τους 5 δείκτες και τη σταθερά.

Coefficient: Η δεύτερη στήλη περιέχει τις υπολογισμένες παραμέτρους, τη λύση δηλαδή της εξίσωσης.

Std. Error: Η τρίτη στήλη αναφέρεται στα εκτιμημένα τυπικά σφάλματα των εκτιμήσεων των παραμέτρων. Τα τυπικά σφάλματα είναι στην ουσία οι τετραγωνικές ρίζες των διαγώνιων στοιχείων του πίνακα συνδιακυμάνσεων των παραμέτρων και μετράνε την στατιστική σημαντικότητα τους. Όσο πιο μεγάλα είναι τόσο πιο μεγάλο και το statistical noise στις εκτιμήσεις τους.

z-Statistic: Η τέταρτη στήλη μας δείχνει το z-statistic των παραμέτρων που δεν είναι τίποτα άλλο από τον λόγο της εκτιμημένης κάθε φορά παραμέτρου προς το τυπικό της σφάλμα και ελέγχει τη μηδενική υπόθεση η παράμετρος να είναι μηδέν. Για να ερμηνεύσουμε όμως το z-statistic θα πρέπει να μελετήσουμε την πιθανότητα να παρατηρήσουμε ένα z-statistic δεδομένου ότι η παράμετρος είναι μηδέν.

Prob.: Η τελευταία στήλη μας δείχνει αυτήν ακριβώς την πιθανότητα που είναι γνωστή και ως p-value ή marginal significance level. Παρατηρώντας την p-value μπορούμε με μια ματιά να δεχτούμε ή να απορρίψουμε την υπόθεση ότι η πραγματική παράμετρος είναι μηδέν. Από τη στιγμή που τα τεστ γίνονται όλα στο επίπεδο στατιστικής σημαντικότητας 0.05, μία τιμή της p-value μικρότερη από αυτό το νούμερο μας επιτρέπει να απορρίψουμε τη μηδενική υπόθεση και να θεωρήσουμε την παράμετρό μας στατιστικά σημαντική. Αυτή ακριβώς η p-value αποτέλεσε και το ένα από τα κριτήρια επιλογής των 5 καλύτερων μεταβλητών. Γι’ αυτό άλλωστε και παρατηρούμε ότι οι συγκεκριμένες 5 μεταβλητές διαθέτουν πολύ καλές p-values όλες μικρότερες κατά πολύ του 0.05. Μόνο η σταθερά δεν παρουσιάζει καλή p-value γι’ αυτό και δεν τη συμπεριλαμβάνουμε στο μοντέλο μας.

5.1.2.Συσχέτιση μεταβλητών:

Πίνακας συνδιακυμάνσεων παραμέτρων						
	C	X ₂	X ₆	X ₉	X ₁₁	X ₁₃
C	0.1247	-0.1453	0.0128	0.0095	-0.0257	0.0037
X ₂	-0.1453	0.7319	-0.5284	-0.0271	-0.2305	-0.1313

X_6	0.0128	-0.5284	1.1074	0.0201	0.2475	0.1882
X_9	0.0095	-0.0271	0.0201	0.0137	0.0029	0.0210
X_{11}	-0.0257	-0.2305	0.2475	0.0029	0.1541	0.0304
X_{13}	0.0037	-0.1313	0.1882	0.0210	0.0304	0.2339

Πίνακας συσχετίσεων						
	C	X_2	X_6	X_9	X_{11}	X_{13}
C	1	-0.481	0.034	0.230	-0.185	0.022
X_2	-0.481	1	-0.587	-0.271	-0.686	-0.317
X_6	0.034	-0.587	1	0.163	0.599	0.370
X_9	0.230	-0.271	0.163	1	0.063	0.371
X_{11}	-0.185	-0.686	0.599	0.063	1	0.160
X_{13}	0.022	-0.317	0.370	0.371	0.160	1

Οι πίνακες συνδιακυμάνσεων και συσχετίσεων των παραμέτρων μας δείχνουν τη συσχέτιση ανάμεσα στη σταθερά και τις μεταβλητές μεταξύ τους και επομένως και την αλληλεπίδρασή τους μέσα στο μοντέλο. Ο πίνακας συσχετίσεων μας δίνει μια καλύτερη πληροφόρηση καθώς η τιμή του συντελεστή συσχέτισης R κυμαίνεται μεταξύ του -1 και του 1 και επομένως μπορούμε με μια ματιά να δούμε πόσο μεγάλη συσχέτιση παρουσιάζουν οι μεταβλητές μεταξύ τους.

Παρατηρούμε ότι οι συσχετίσεις ανάμεσα στις μεταβλητές που επιλέξαμε δεν είναι πάρα πολύ μεγάλες, με την έννοια ότι καμία δεν πλησιάζει το 1 . Η μεγαλύτερη που εμφανίζεται είναι της τάξεως του 0.68 . Παρ' όλ' αυτά, οι συσχετίσεις των υπόλοιπων μεταβλητών που αποκλείστηκαν από το μοντέλο, ήταν μικρότερες από αυτές. Οι υπόλοιποι δείκτες δεν εμφανίζουν ιδιαίτερη συσχέτιση και αυτό, μαζί και με άλλα στατιστικά κριτήρια που αναφέρονται παραπάνω, καθώς και στον ΠΙΝΑΚΑ 1 του παραρτήματος, αποτελούν τις αιτίες για τις οποίες δεν συμπεριλήφθηκαν στο μοντέλο.

Αν πάρουμε λοιπόν μία μία τις συσχετίσεις καθεμιάς μεταβλητής με όλες τις υπόλοιπες, παρατηρούμε καταρχήν την αρνητική συσχέτιση του x_2 με όλες τις υπόλοιπες. Η συσχέτιση αυτή δικαιολογεί και το αρνητικό πρόσημο του x_2 μέσα στην εξίσωση. Ο δείκτης αυτός (Κυκλοφορούν Ενεργητικό / Ενεργητικό) δεν αναμενόταν να παρουσιάζει αρνητικό πρόσημο, καθώς μία μείωση του δείκτη συνεπάγεται χειροτέρευση της εταιρίας ενώ αύξησή του αποτελεί θετικό σημάδι, όπως άλλωστε συμβαίνει και με όλους τους υπόλοιπους δείκτες. Θεωρητικά δηλαδή, μία αύξηση του Κυκλοφορούντος Ενεργητικού, δηλαδή των στοιχείων του Ενεργητικού που η επιχείρηση μπορεί να ρευστοποιήσει, έχει ως αποτέλεσμα την βελτίωση της κατάστασής της, καθώς μπορεί ανά πάσα στιγμή να ρευστοποιήσει αυτά τα στοιχεία για να ξεπληρώσει τις υποχρεώσεις της. Όσο μεγαλύτερο Κυκλοφορούν Ενεργητικό έχει δηλαδή μια εταιρία τόσο πιο υγιής θεωρείται, καθώς είναι καλυμμένη σε κάθε περίπτωση άμεσου χρέους που μπορεί να προκύψει.

Αυξημένος δηλαδή δείκτης Κυκλοφορούντος Ενεργητικού / Ενεργητικό συνδυάζεται με υγιή επιχείρηση. Αυτό ισχύει θεωρητικά. Πρακτικά όμως, μπορεί να συμβαίνει και το αντίθετο. Δηλαδή ένας μειωμένος τέτοιος δείκτης να συνεπάγεται υγιή επιχείρηση. Αυτό μπορεί να συμβαίνει γιατί το μειωμένο Κυκλοφορούν Ενεργητικό της επιχείρησης μπορεί να προέρχεται από μειωμένα εμπορεύματα. Η εταιρία δηλαδή μπορεί να έχει μειωμένο Κυκλοφορούν Ενεργητικό γιατί πουλάει συνέχεια τα εμπορεύματά της και τα προϊόντα της δεν μένουν καθόλου στην επιχείρηση. Κάτι τέτοιο όμως σημαίνει ότι η επιχείρηση “πάει καλά” και είναι μάλλον υγιής παρά προβληματική. Όλα εξαρτώνται δηλαδή από την κυκλοφοριακή ταχύτητα του Κυκλοφορούντος Ενεργητικού της επιχείρησης. Μία υγιής επιχείρηση με άλλα λόγια με μεγάλη κυκλοφοριακή ταχύτητα Κυκλοφορούντος Ενεργητικού μπορεί να έχει κάλλιστα έναν μικρό δείκτη Κυκλοφορούντος Ενεργητικού / Ενεργητικό. Κάτι τέτοιο μπορεί να συμβαίνει και με τις εταιρίες του συγκεκριμένου δείγματος που χρησιμοποιήσαμε, καθώς μέσα σε αυτές περιέχονται για παράδειγμα εταιρίες παραγωγής και πώλησης ενδυμάτων όπου, τα προϊόντα μπορεί πράγματι να πωλούνται διαρκώς και με μεγάλη ταχύτητα. Αυτή είναι η οικονομική εξήγηση που μπορούμε να δώσουμε στην αντίθετη αυτή συμπεριφορά του δείκτη.

Από στατιστική άποψη τώρα, η συμπεριφορά αυτή μπορεί να ερμηνευθεί βάσει της αλληλοσυσχέτισης των μεταβλητών. Το αρνητικό πρόσημο του χ_2 θα έχει ως αποτέλεσμα μία αύξησή του να επιδεινώσει την κατάσταση της επιχείρησης και μία μείωσή του να τη βελτιώσει. Η συμπεριφορά αυτή είναι όπως είπαμε αντίθετη από τη λογική αναμενόμενη. Η τελευταία όμως, χαρακτηρίζει τη μεταβλητή όταν επιδρά μόνη της στο y . Επειδή ακριβώς κάθε μεταβλητή ασκεί μία επίδραση στο y σε συνεργασία με όλες τις υπόλοιπες μεταβλητές και όχι μόνη της, μια απόπειρα να εξηγήσουμε τα εκτιμημένα πρόσημα σύμφωνα με τις επιμέρους επιδράσεις θα ήταν παραπλανητική. Σε συνεργασία λοιπόν με όλες τις υπόλοιπες μεταβλητές και εξαιτίας της συσχέτισής της με αυτές, η χ_2 εμφανίζει μια τελείως διαφορετική συμπεριφορά. Απόδειξη αποτελεί αυτός ακριβώς ο πίνακας των συσχετίσεων των παραμέτρων που μας φανερώνει την αρνητική συσχέτιση του συγκεκριμένου δείκτη με τους υπόλοιπους και επομένως και την αιτιολογία του αρνητικού του προσήμου. Επειδή λοιπόν τα συγκεκριμένα μοντέλα αυτό το νόημα έχουν, να προβλέψουν μια κατάσταση με τη βοήθεια και τη συνεργασία πολλών μεταβλητών μαζί, δεν θα ήταν εύλογο να περιοριστούμε στις επιμέρους επιδράσεις των τελευταίων.

Όλοι οι άλλοι δείκτες παρουσιάζουν θετική συσχέτιση μεταξύ τους γι’ αυτό και έχουν θετικά πρόσημα οι παράμετροί τους μέσα στην εξίσωση. Αυτό σημαίνει ότι η αύξηση κάποιου από αυτούς θα επιδράσει θετικά στο y και θα συνοδεύεται από μία αύξηση και των υπολοίπων. Αυτό είναι απόλυτα λογικό, καθώς μία εταιρία για παράδειγμα με αυξημένο δείκτη Κέρδη προ φόρων / Βραχυπρόθεσμες υποχρεώσεις, θα αναμένουμε να εμφανίζει επίσης αυξημένο δείκτη Πωλήσεις / Ενεργητικό.

Αυτό που αξίζει να παρατηρήσουμε είναι ότι τα λογιστικά στοιχεία που βοηθούν τελικά στην πρόβλεψη του πιστωτικού κινδύνου για μια επιχείρηση περιορίζονται σε στοιχεία κερδοφορίας, ρευστότητας και αποδοτικότητας. Πρόκειται για στοιχεία όπως το Ενεργητικό της επιχείρησης και υποκατηγορίες του, που εκφράζουν την περιουσιακή κατάσταση της επιχείρησης καθώς και τη δυνατότητά της να ρευστοποιήσει στοιχεία για να εξοφλήσει τις υποχρεώσεις της, το Παθητικό της και

υποκατηγορίες του που εκφράζουν τις βραχυπρόθεσμες και μακροπρόθεσμες υποχρεώσεις της εταιρίας, στοιχείο καίριο για την αξιολόγηση της κατάστασής της, τα ίδια και ξένα κεφάλαια που διαθέτει η επιχείρηση ως εγγύηση και κάλυψη των χρεών και τέλος τα κέρδη και οι πωλήσεις που επίσης προσδιορίζουν την ευρωστία της επιχείρησης και τη δυνατότητα κάλυψης των χρεών της με αυτά.

Λογιστικά, αυτοί είναι οι συνηθέστεροι αριθμοδείκτες που χρησιμοποιούνται στις περισσότερες πιστωτικές αναλύσεις εταιριών, κάτι που επιβεβαιώσαμε και με το συγκεκριμένο μοντέλο.

5.1.3. Στατιστική σημαντικότητα μοντέλου- Τεστ:

Αρχικό δείγμα (1yr prior)

Πίνακας στατιστικών μέτρων μοντέλου			
Mean dependent var	0.5	S.D. dependent var	0.503003
S.E. of regression	0.397119	McFadden R-squared	0.369757
Log likelihood	-36.69548		
LR statistic (5 df)	43.05777		
Probability (LR stat)	3.60E-08		

Ο παραπάνω πίνακας παρουσιάζει στατιστικά αποτελέσματα που προσδιορίζουν το μοντέλο ως σύνολο:

Mean dependent var: παρατηρούμε τον μέσο της εξαρτημένης μεταβλητής y που κυμαίνεται στο 0.5, νούμερο που δικαιολογείται από το γεγονός ότι οι τιμές που έχουμε δώσει για το y ως δεδομένα στο πρόγραμμα είναι 0 και 1.

S.D. dependent var: μας δείχνει την τυπική απόκλιση του y που κυμαίνεται και αυτή στο 0.5, αριθμός που επίσης δικαιολογείται από τις τιμές που έχουμε δώσει σε αυτό.

S.E. of regression: υπολογίζεται με βάση τη διακύμανση των καταλοίπων και μας δίνει το τυπικό σφάλμα της παλινδρόμησης, που δεν εμφανίζεται ιδιαίτερα αυξημένο.

Log likelihood: μας δίνει τη μέγιστη τιμή της συνάρτησης μέγιστης πιθανοφάνειας.

LR statistic: το μέτρο αυτό ελέγχει την από κοινού υπόθεση ότι όλες οι παράμετροι εκτός της σταθεράς είναι μηδενικές. Είναι ανάλογο του F-statistic και ελέγχει τη συνολική στατιστική σημαντικότητα του μοντέλου. (Ο αριθμός στην παρένθεση δείχνει τους βαθμούς ελευθερίας.) Για να το ερμηνεύσουμε πρέπει να κοιτάξουμε το παρακάτω μέτρο που είναι η πιθανότητά του.

Probability (LR stat): είναι η p -value του LR test statistic. Όταν είναι μικρότερη του 0.05 μπορούμε να θεωρήσουμε ότι όλες οι εκτιμημένες παράμετροι είναι διάφορες του μηδενός και άρα το μοντέλο μας είναι στατιστικά σημαντικό. Παρατηρούμε πως εδώ είναι ένα νούμερο πολύ κοντά στο μηδέν.

McFadden R-squared: είναι ένας δείκτης του likelihood ratio και είναι ανάλογο του R^2 , δείχνει δηλαδή το ποσοστό ερμηνείας που μας δίνουν οι ανεξάρτητες μεταβλητές για το y . Έχει την ιδιότητα να κυμαίνεται μεταξύ του 0 και του 1, δείχνοντας μεγαλύτερη αποτελεσματικότητα του μοντέλου καθώς πλησιάζει το 1. Εδώ, καθώς και σε όλες τις δοκιμές που κάναμε με τους δείκτες, κυμαίνεται στο 0.36.

5.1.4.Ακρίβεια πρόβλεψης μοντέλου:

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Probit			
Date: 05/11/02 Time: 15:32			
Sample (adjusted): 1 84			
Included observations: 84 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P			
(Dep=1)<=C	33	10	43
P (Dep=1)>C	9	32	41
Total	42	42	84
Correct	33	32	65
% Correct	78.57	76.19	77.38
% Incorrect	21.43	23.81	22.62

Ο παραπάνω πίνακας παρουσιάζει την ακρίβεια με την οποία προβλέπει το μοντέλο μας. Η ακρίβεια αναμένεται σχετικά αυξημένη καθώς το τεστ που διεξάγει το πρόγραμμα γίνεται επί του αρχικού δείγματος βάσει του οποίου δημιουργήθηκε το μοντέλο. Πράγματι, παρατηρούμε ότι η ακρίβεια είναι ικανοποιητική, της τάξεως του 77.38% για το σύνολο του μοντέλου. Το cutoff score βάσει του οποίου εκτιμήθηκε η συγκεκριμένη ακρίβεια είναι 0.5 και αναγράφεται στο επάνω μέρος του πίνακα. Αυτό σημαίνει ότι αν οι εκτιμημένες τιμές του y ήταν πάνω από 0.5 το πρόγραμμα τις λάμβανε ίσες με τη μονάδα ενώ αν ήταν κάτω από 0.5 τις λάμβανε ως μηδενικές. Μία εταιρία δηλαδή που θα παρουσίαζε σκορ 0.55 θα θεωρούταν από το πρόγραμμα ως υγιής. Παρακάτω διεξάγουμε το ίδιο τεστ με διαφορετικά cutoff scores για να μελετήσουμε πώς εξελίσσεται η ακρίβεια του μοντέλου. Αυτό

που αξίζει να επισημάνουμε είναι ότι το ποσοστό ακρίβειας στην πρόβλεψη των καθυστερημένων εταιριών παρουσιάζεται μεγαλύτερο από το αντίστοιχο στην πρόβλεψη υγιών εταιριών, μειώνοντας έτσι το Λάθος τύπου I, κάτι που είναι αρκετά σημαντικό. Γιατί μας ενδιαφέρει περισσότερο το μοντέλο μας να προβλέπει ακριβέστερα τις καθυστερημένες εταιρίες, καθώς από αυτές θα προκληθούν απώλειες αν δεν γίνει σωστή πρόβλεψη. Είναι με άλλα λόγια προτιμότερο να κάνουμε σφάλμα τύπου II, να προβλέψουμε προβληματική μια υγιή εταιρία, το πολύ-πολύ να μην τη χρηματοδοτήσουμε, παρά να κάνουμε σφάλμα τύπου I, να προβλέψουμε δηλαδή ως υγιή μια εταιρία προβληματική η οποία δεν θα μας επιστρέψει ποτέ τη δανειοδότηση που θα της δώσουμε.

Στην πρώτη στήλη του πίνακα βλέπουμε ότι 33 από τις 42 προβληματικές εταιρίες προβλέφθηκαν σωστά και 9 λάθος, δίνοντάς μας ένα ποσοστό 78.57% σωστής πρόβλεψης και 21.43% λανθασμένης. Η δεύτερη στήλη μας δίνει τα αντίστοιχα νούμερα και ποσοστά για τις υγιείς εταιρίες. Η τρίτη στήλη μας δείχνει τα ίδια αποτελέσματα επί του συνόλου των παρατηρήσεων.

Παρακάτω παρατίθενται τα αποτελέσματα του ίδιου τεστ με μικρές μεταβολές του cutoff score:

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Probit			
Date: 05/11/02 Time: 15:32			
Sample (adjusted): 1 84			
Included observations: 84 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)≤C	29	6	35
P (Dep=1)>C	13	36	49
Total	42	42	84
Correct	29	36	65
% Correct	69.05	85.71	77.38
% Incorrect	30.95	14.29	22.62

Παρατηρούμε ότι η γενική ακρίβεια του μοντέλου δεν μεταβάλλεται, μεταβάλλονται όμως οι επιμέρους. Συγκεκριμένα, η ακρίβεια πρόβλεψης των προβληματικών μειώνεται, ενώ των υγιών αυξάνεται, κι αυτό γιατί μειώνοντας το σκορ, το πρόγραμμα λαμβάνει περισσότερες εταιρίες πλέον ως υγιείς, δίνοντας έτσι μεγαλύτερο περιθώριο διόρθωσης σ' αυτόν τον τομέα και μικρότερο περιθώριο λάθους στον άλλον. Εταιρίες με άλλα λόγια που βρίσκονταν στο όριο του 0.5 και θεωρήθηκαν προηγουμένως ως προβληματικές αλλά ήταν υγιείς λαμβάνονται τώρα πλέον ως υγιείς, αλλά εταιρίες που ήταν προβληματικές και που

με το προηγούμενο σκορ προβλέπονταν σωστά τώρα λαμβάνονται επίσης ως υγιείς, όποτε προβλέπονται λάθος.

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Probit			
Date: 05/11/02 Time: 15:32			
Sample (adjusted): 1 84			
Included observations: 84 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	37	13	50
P (Dep=1)>C	5	29	34
Total	42	42	84
Correct	37	29	66
% Correct	88.1	69.05	78.57
% Incorrect	11.9	30.95	21.43

Η αύξηση του cutoff point παρατηρούμε ότι αυξάνει τη γενική ακρίβεια του μοντέλου, κάτι που δικαιολογείται από τη μεταβολή στις επιμέρους. Προφανώς ένα πιο αυξημένο σκορ είναι και πιο σωστό καθώς δίνει την ευκαιρία στο πρόγραμμα να συμπεριλάβει περισσότερες προβληματικές εταιρίες στη σωστή του πρόβλεψη, μειώνοντας έτσι το σφάλμα τύπου I. Υπήρχαν επομένως και εταιρίες με σκορ 0.5 και λίγο παραπάνω που παρ' όλ' αυτά ήταν προβληματικές.

5.1.5. Goodness of fit test:

Το τεστ αυτό συγκρίνει τις αναμενόμενες τιμές με τις πραγματικές ανά γκρουπ. Αν οι διαφορές που θα προκύψουν είναι μεγάλες, απορρίπτουμε το μοντέλο ως μη αντιπροσωπευτικό των αληθινών δεδομένων. Το Hosmer-Lemeshow τεστ που διεξήγαμε μας έδωσε μια χαμηλή τιμή της τάξεως του H-L Statistic: 3.4785, που δεν μας επιτρέπει να απορρίψουμε το μοντέλο μας.

Στη συνέχεια παραθέτουμε μία σειρά από τεστ για την επιβεβαίωση της ακρίβειας του μοντέλου μας :

Τεστ 1: Αρχικό δείγμα 2yrs prior

Χρησιμοποιώντας ως βάση τα δεδομένα του αρχικού πάλι δείγματος αλλά για 2 έτη πριν την καθυστέρηση, εξετάσαμε την ακρίβεια πρόβλεψης του μοντέλου. Εισάγαμε δηλαδή ως δεδομένα τους 5 δείκτες των 54 πια εταιριών, τις παραμέτρους του μοντέλου μας που είναι πια γνωστές και επιλύσαμε ως προς το άγνωστο πλέον y . Στη συνέχεια, συγκρίναμε τα εκτιμημένα αυτά y με τα πραγματικά και διαπιστώσαμε το ποσοστό ακρίβειας του μοντέλου μας σε αυτό το δείγμα. Παρακάτω παρατίθεται ο πίνακας των αποτελεσμάτων με διαφορετικό κάθε φορά cutoff score:

test1			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 54			
Included observations: 54 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)<=C	23	8	31
P (Dep=1)>C	4	19	23
Total	27	27	54
Correct	23	19	42
% Correct	85.19	70.37	77.78
% Incorrect	14.81	29.63	22.22

Test1			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 54			
Included observations: 54 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)<=C	23	6	29
P (Dep=1)>C	4	21	25
Total	27	27	54
Correct	23	21	44
% Correct	85.19	77.78	81.48
% Incorrect	14.81	22.22	18.52

Test1			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 54			
Included observations: 54 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)≤C	23	10	33
P (Dep=1)>C	4	17	21
Total	27	27	54
Correct	23	17	40
% Correct	85.19	62.96	74.07
% Incorrect	14.81	37.04	25.93

Παρατηρούμε ότι η συνολική ακρίβεια του μοντέλου μας βρίσκεται και πάλι στο ικανοποιητικό επίπεδο του 77.78%, ελάχιστα παραπάνω από την ακρίβεια στο αρχικό δείγμα από το οποίο δημιουργήθηκε το μοντέλο. Βέβαια μην ξεχνάμε ότι και αυτό το τεστ έγινε σε ένα μικρότερο που προήλθε όμως από το ίδιο αρχικό δείγμα αλλά για δύο έτη πριν την καθυστέρηση. Πάλι παρατηρούμε μεγαλύτερη ακρίβεια στην πρόβλεψη των καθυστερημένων.

Στη συνέχεια, παρατηρούμε ότι η μείωση του cutoff score έχει ως συνέπεια να αυξηθεί η συνολική ακρίβεια του μοντέλου αφήνοντας αμετάβλητη την ακρίβεια στις προβληματικές αλλά αυξάνοντας την αντίστοιχη στις υγιείς, για τους ίδιους λόγους που παραθέσαμε και νωρίτερα. Η αύξηση από την άλλη του σκορ μειώνει την προβλεψιμότητα του μοντέλου αφήνοντας για μια ακόμη φορά αμετάβλητη την ακρίβεια στις προβληματικές αλλά μειώνοντας εκείνη των υγιών. Επομένως, αυτό που διαφοροποιείται στην περίπτωση αυτή από την αυξομείωση του σκορ είναι μόνο η ακρίβεια ουσιαστικά στην πρόβλεψη των υγιών. Σε αντίθεση με το προηγούμενο τεστ, που έγινε στο ίδιο ακριβώς αρχικό δείγμα από το οποίο προήλθε το μοντέλο και όπου το 0.6 εμφανίστηκε ως καλύτερο σκορ, εδώ παρατηρούμε ότι το 0.4 παρουσιάζεται ως καλύτερο σκορ.

Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior

Στη συνέχεια κάναμε το ίδιο ακριβώς τεστ στα στοιχεία όλων των εταιριών του αρχικού δείγματος και για το 1 και για τα 2 έτη πριν τη χρεωκοπία. Τα αποτελέσματα παρατίθενται πάλι για διαφορετικά cutoff scores:

Test2			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 138			
Included observations: 138 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	60	22	82
P (Dep=1)>C	9	47	56
Total	69	69	138
Correct	60	47	107
% Correct	86.96	68.12	77.54
% Incorrect	13.04	31.88	22.46

Test2			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 138			
Included observations: 138 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	60	19	79
P (Dep=1)>C	9	50	59
Total	69	69	138
Correct	60	50	110
% Correct	86.96	72.46	79.71
% Incorrect	13.04	27.54	20.29

Test2			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 138			
Included observations: 138 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)<=C	61	26	87
P (Dep=1)>C	8	43	51
Total	69	69	138
Correct	61	43	104
% Correct	88.41	62.32	75.36
% Incorrect	11.59	37.68	24.64

Παρατηρούμε ότι η γενική ακρίβεια με cutoff score 0.5 παραμένει στο ίδιο περίπου επίπεδο με ελαφρώς αυξημένη την ακρίβεια στις προβληματικές και ελαφρώς μειωμένη την αντίστοιχη στις υγιείς. Μειώνοντας το cutoff score αυξάνεται πάλι η συνολική ακρίβεια του μοντέλου επηρεαζόμενη ξανά από την αντίστοιχη μεταβολή μόνο της ακρίβειας στις υγιείς, αφού αυτή των προβληματικών παραμένει κι εδώ αμετάβλητη. Η αύξηση αντίθετα του σκορ έχει ως αποτέλεσμα τη μείωση της συνολικής ακρίβειας κάτι που προέρχεται για μια ακόμη φορά από την επίδραση της ακρίβειας των υγιών, καθώς φαίνεται να επισκιάζει τελείως την οποιαδήποτε αντίστοιχη επίδραση της μικρής αύξησης της προβλεψιμότητας στις προβληματικές. Για μια ακόμη φορά το 0.4 αποδεικνύεται καλύτερο cutoff score.

Τεστ 3: 2^ο δείγμα 2yrs prior

Αυτό όμως που έχει ιδιαίτερο ενδιαφέρον και θα μας διαφωτίσει περισσότερο για την αποτελεσματικότητα του μοντέλου μας είναι ένας έλεγχος με δείγμα διαφορετικό από το πρωτογενές μας. Πράγματι διεξήγαμε και τέτοιους ελέγχους με τη βοήθεια του δεύτερου δείγματος που δημιουργήσαμε. Πρώτα διεξήγαμε έναν έλεγχο στις παρατηρήσεις αυτού του δείγματος για 2 έτη πριν την καθυστέρηση. Δυστυχώς οι παρατηρήσεις αυτές δεν ήταν αρκετές ώστε το δείγμα μας να θεωρηθεί

αντιπροσωπευτικό. Παρ' όλ' αυτά παραθέτουμε τα αποτελέσματα με διαφορετικά πάντα cutoff scores:

Test3			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	9	4	13
P (Dep=1)>C	0	5	5
Total	9	9	18
Correct	9	5	14
% Correct	100.00	55.56	77.78
% Incorrect	0.00	44.44	22.22

Test3			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	9	4	13
P (Dep=1)>C	0	5	5
Total	9	9	18
Correct	9	5	14
% Correct	100.00	55.56	77.78
% Incorrect	0.00	44.44	22.22

Test3			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)≤C	9	4	13
P (Dep=1)>C	0	5	5
Total	9	9	18
Correct	9	5	14
% Correct	100.00	55.56	77.78
% Incorrect	0.00	44.44	22.22

Παρατηρούμε ότι οι παρατηρήσεις για τις οποίες έγινε ο παραπάνω έλεγχος ήταν μόνο 18 και γι' αυτό ακριβώς συναντάμε τα παράδοξα αποτελέσματα της ακρίβειας 100% στις προβληματικές εταιρίες (στις 18 μόνο παρατηρήσεις έτυχε να μη γίνει καμία λάθος πρόβλεψη για προβληματική εταιρία), καθώς και την επιμονή στα ίδια ακριβώς αποτελέσματα παρά την αλλαγή του cutoff point. Η συνολική ακρίβεια του μοντέλου παρουσιάζεται κι εδώ στο επίπεδο του 77.78%.

Τεστ 4: 2^ο δείγμα 1yr prior

Ο πιο ενδιαφέρων και αντιπροσωπευτικός ίσως έλεγχος είναι αυτός που έγινε στο δεύτερο αυτό δείγμα για 1 έτος πριν την καθυστέρηση. Το δείγμα αυτό περιείχε 38 παρατηρήσεις και μας έδωσε τα παρακάτω αποτελέσματα:

Test4			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	17	7	24
P (Dep=1)>C	2	12	14
Total	19	19	38
Correct	17	12	29
% Correct	89.47	63.16	76.32
% Incorrect	10.53	36.84	23.68

Test4			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	17	6	23
P (Dep=1)>C	2	13	15
Total	19	19	38
Correct	17	13	30
% Correct	89.47	68.42	78.95
% Incorrect	10.53	31.58	21.05

Test4			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)<=C	17	8	25
P (Dep=1)>C	2	11	13
Total	19	19	38
Correct	17	11	28
% Correct	89.47	57.89	73.68
% Incorrect	10.53	42.11	26.32

Παρατηρούμε μια συνολική ακρίβεια της τάξεως του 76.32%, με την ακρίβεια στην πρόβλεψη των καθυστερημένων να υπερέχει αισθητά εκείνης των υγιών. Η αυξομείωση του cutoff point αφήνει αμετάβλητη την ακρίβεια στις προβληματικές και αυξομειώνει αντίστροφα πάντα την ακρίβεια στις υγιείς και τη συνολική ακρίβεια. Πάλι το 0.4 αποδεικνύεται ως καλύτερο cutoff point. Ο έλεγχος αυτός αποδεικνύει ότι το μοντέλο προβλέπει και κατατάσσει σωστά και με ικανοποιητική ακρίβεια και άλλες παρατηρήσεις – εταιρίες εκτός αυτών του αρχικού δείγματος από το οποίο προήλθε το μοντέλο.

Τεστ 5: Αρχικό δείγμα (1yr prior)- ΕΠΑΝΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

Τέλος, επειδή ένα σετ μεταβλητών μπορεί να είναι αποτελεσματικό για το αρχικό δείγμα αλλά όχι και για τον πληθυσμό, εξετάσαμε την robustness του μοντέλου, την ικανότητά του δηλαδή να προβλέπει σωστά και για παρατηρήσεις άλλες εκτός των αρχικών. διεξήγαμε λοιπόν ένα validation test ως εξής: επανεκτιμήσαμε τις παραμέτρους του μοντέλου κρατώντας σταθερές τις 5 μεταβλητές μας αλλά χρησιμοποιώντας αυτή τη φορά ως δείγμα ένα υποσύνολο του αρχικού δείγματος από το οποίο προήλθε το μοντέλο (42 πρώτες παρατηρήσεις 1 έτους πριν την καθυστέρηση). Η νέα εξίσωση που προέκυψε είναι η εξής:

$$y = - 5.414 \cdot X_2 + 10.141 \cdot X_6 + 2.396 \cdot X_9 + 2.097 \cdot X_{11} + 2.747 \cdot X_{13}$$

Στη συνέχεια, ελέγξαμε την ακρίβεια αυτού του μοντέλου με τη βοήθεια του υπόλοιπου δείγματος. Χρησιμοποιήσαμε δηλαδή ως δεδομένα τις νέες παραμέτρους, τους 5 δείκτες των υπόλοιπων 42 εταιριών για 1 έτος πριν την καθυστέρηση και λύσαμε ως προς y.

Συγκρίναμε τέλος τα εκτιμημένα αυτά y με τα πραγματικά που ήδη γνωρίζαμε και βρήκαμε την ακρίβεια του μοντέλου:

Test5			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)≤C	18	11	29
P (Dep=1)>C	3	10	13
Total	21	21	42
Correct	18	10	28
% Correct	85.71	47.62	66.67
% Incorrect	14.29	52.38	33.33

Test5			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)≤C	19	11	30
P (Dep=1)>C	2	10	12
Total	21	21	42
Correct	19	10	29
% Correct	90.48	47.62	69.05
% Incorrect	9.52	52.38	30.95

Test5			
Dependent Variable: Y			
Method: ML - Binary Probit			
Sample (adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P (Dep=1)≤C	19	11	30
P (Dep=1)>C	2	10	12
Total	21	21	42
Correct	19	10	29
% Correct	90.48	47.62	69.05
% Incorrect	9.52	52.38	30.95

Παρατηρούμε εδώ μια μείωση της ακρίβειας του μοντέλου με την εξής όμως ιδιομορφία: η ακρίβεια πρόβλεψης για τις καθυστερημένες επιχειρήσεις έχει αυξηθεί εντυπωσιακά, ενώ αντίθετα για τις υγιείς έχει μειωθεί δραματικά. Μία ακόμη ιδιομορφία παρουσιάζεται όταν αυξομειώνουμε το cutoff score όπου παρατηρούμε ακριβώς τις ίδιες αλλαγές είτε το μειώνουμε είτε το αυξάνουμε: τη συνολική αύξηση της ακρίβειας η οποία προέρχεται αποκλειστικά και μόνο από την αντίστοιχη αύξηση στην ακρίβεια πρόβλεψης των προβληματικών. Η αντίστοιχη ακρίβεια στις υγιείς παραμένει αμετάβλητη. Οι ιδιομορφίες αυτές, καθώς και η συνολική μείωση της ακρίβειας του μοντέλου ίσως να οφείλονται στο ότι οι καινούργιες παράμετροι που εκτιμήθηκαν προήλθαν από μικρότερο και γι' αυτό λιγότερο αποτελεσματικό δείγμα.

5.1.6.Γενικά συμπεράσματα για το probit:

Τα γενικά πάντως συμπεράσματα που μπορούμε να βγάλουμε από τα παραπάνω αποτελέσματα, είναι ότι παρά την ανεπάρκεια αρχικών στοιχείων, το μοντέλο probit που καταφέραμε να δημιουργήσουμε παρέχει ικανοποιητική ακρίβεια στην πρόβλεψη και κατάταξη των εταιριών σε υγιείς και προβληματικές, υπερτερώντας αισθητά στην πρόβλεψη των προβληματικών. Ακόμη και αν θεωρήσουμε ότι η ακρίβειά του περιορίζεται τελικά στο 67% του τελευταίου τεστ, αν και τονίσαμε ότι το συγκεκριμένο τεστ δεν είναι απαραίτητα αντιπροσωπευτικό, εξακολουθεί να είναι ικανοποιητική λαμβάνοντας υπόψη τα δεδομένα που είχαμε στη διάθεσή μας. Ως αντιπροσωπευτικότερο cutoff score πρέπει μάλλον να θεωρήσουμε το 0.4, αφού σε όλα σχεδόν τα τεστ μας δίνει μεγαλύτερη ακρίβεια. Επιπλέον, ο περιορισμός του σφάλματος τύπου I που προσφέρει το συγκεκριμένο μοντέλο μας ενδιαφέρει ιδιαίτερα στο συγκεκριμένο πρόβλημα πιστωτικού κινδύνου και γι' αυτό το λόγο το θεωρούμε ως σημαντικό πλεονέκτημα του μοντέλου.

5.2.Logit

5.2.1.Στατιστική σημαντικότητα μεταβλητών:

Πίνακας στατιστικών μέτρων μεταβλητών				
Dependent Variable: Y				
Method: ML - Binary Logit				
Date: 05/11/02 Time: 20:27				
Sample (adjusted): 1 84				
Included observations: 84 after adjusting endpoints				
Convergence achieved after 5 iterations				
Covariance matrix computed using second derivatives				
Variable	Coefficient	Std. Error	z-Statistic	Prob.
C	-0.34123	0.617466	-0.55262	0.5805
X ₂	-4.38228	1.524241	-2.87506	0.004
X ₆	7.462822	1.886104	3.956739	0.0001
X ₁₀	0.432448	0.195646	2.210366	0.0271
X ₁₁	2.521026	0.762687	3.305455	0.0009
X ₁₃	2.382978	0.820655	2.90375	0.0037

Τα μέτρα που εμφανίζονται στον παραπάνω πίνακα έχουν ήδη αναλυθεί κατά την επεξήγηση των αποτελεσμάτων του probit. Αυτό που μας ενδιαφέρει κυρίως από αυτόν τον πίνακα είναι οι εκτιμήσεις των παραμέτρων των μεταβλητών και της σταθεράς, που μας δίνουν στην ουσία το μοντέλο μας, καθώς και οι p-values τους. Παρατηρούμε ότι οι τελευταίες εμφανίζονται πολύ μικρότερες του 0.05, κάτι που μας επιτρέπει να δεχτούμε τις μεταβλητές μας ως στατιστικά σημαντικές και να τις εντάξουμε στο μοντέλο μας. Το ίδιο όμως δεν ισχύει και για τον όρο της σταθεράς, ο οποίος εμφανίζει $p > 0.05$ και γι' αυτό το λόγο δεν τον εντάσσουμε καθόλου στο μοντέλο. Όσο για τα τυπικά σφάλματα των μεταβλητών εμφανίζονται επίσης αρκετά μικρά ώστε να ανησυχούμε για statistical noise στις εκτιμήσεις τους.

5.2.2.Συσχέτιση μεταβλητών:

	Πίνακας συνδιακυμάνσεων παραμέτρων					
	C	X ₂	X ₆	X ₁₀	X ₁₁	X ₁₃
C	0.3813	-0.4089	0.0153	0.0258	-0.0996	-0.0015
X ₂	-0.4089	2.3233	-1.8008	-0.0869	-0.8274	-0.4244
X ₆	0.0153	-1.8008	3.5574	0.0813	0.8972	0.6585
X ₁₀	0.0258	-0.0869	0.0813	0.0383	0.0156	0.0655
X ₁₁	-0.0996	-0.8274	0.8972	0.0156	0.5817	0.1283
X ₁₃	-0.0015	-0.4244	0.6585	0.0655	0.1283	0.6735

	Πίνακας συσχετίσεων					
	C	X ₂	X ₆	X ₁₀	X ₁₁	X ₁₃
C	1	-0.434	0.013	0.213	-0.211	-0.003
X ₂	-0.434	1	-0.626	-0.291	-0.712	-0.339
X ₆	0.013	-0.626	1	0.220	0.624	0.425
X ₁₀	0.213	-0.291	0.220	1	0.105	0.408
X ₁₁	-0.211	-0.712	0.624	0.105	1	0.205
X ₁₃	-0.003	-0.339	0.425	0.408	0.205	1

Από τον πίνακα των συσχετίσεων, παρατηρούμε ότι οι μεταβλητές παρουσιάζουν ελαφρώς μεγαλύτερη συσχέτιση μεταξύ τους από ότι στο probit, γεγονός που δίνει ένα μεγαλύτερο πλεονέκτημα στο logit, καθώς θέλουμε η συσχέτιση των μεταβλητών μέσα σε κάθε γκρουπ να είναι όσο το δυνατόν μεγαλύτερη ώστε να περιγράφουν και να αντιπροσωπεύουν το κάθε γκρουπ όσο το δυνατόν περισσότερο.

Και εδώ, όπως και στο probit, η μεταβλητή x_2 παρουσιάζει αρνητική συσχέτιση με όλες τις υπόλοιπες μεταβλητές, κάτι που πάλι δικαιολογεί το αρνητικό της πρόσημο μέσα στο μοντέλο. Όλες οι υπόλοιπες μεταβλητές έχουν θετική συσχέτιση μεταξύ τους. Η αρνητική αυτή συσχέτιση της x_2 σημαίνει ότι μια μείωσή της πρέπει να ακολουθηθεί από μία αύξηση των υπολοίπων και το αντίθετο. Σε αντίθεση λοιπόν με τη λογιστική λογική που υπαγορεύει ότι όσο αυξάνεται ο δείκτης Κυκλοφορούν Ενεργητικό / Ενεργητικό μιας επιχείρησης τόσο καλύτερη η κατάστασή της, καθώς θα διαθέτει ρευστότητα, ρευστοποιήσιμα στοιχεία ενεργητικού για να καλύψει τις άμεσες υποχρεώσεις της αν χρειαστεί, στο συγκεκριμένο μοντέλο, ο δείκτης αυτός εμφανίζει μια εντελώς αντίστροφη συμπεριφορά. Το αρνητικό του πρόσημο θα έχει ως αποτέλεσμα μία αύξησή του να επιδεινώσει την κατάσταση της επιχείρησης και μία μείωσή του να τη βελτιώσει.

Παρόλο που αυτό δεν ευσταθεί λογικά, το αποδεχόμαστε με βάση τη λογιστική εξήγηση που δόθηκε και για το probit καθώς και σε όρους συσχέτισης πλέον με τις υπόλοιπες μεταβλητές. Η παραπάνω συμπεριφορά που περιγράψαμε χαρακτηρίζει τη μεταβλητή όταν επιδρά μόνη της στο y . Σε συνεργασία όμως με όλες τις υπόλοιπες και εξαιτίας της συσχέτισής της με αυτές, εμφανίζει μια τελείως διαφορετική συμπεριφορά. Επειδή ακριβώς τα συγκεκριμένα μοντέλα αυτό το νόημα έχουν, να προβλέψουν μια κατάσταση με τη βοήθεια και τη συνεργασία πολλών μεταβλητών μαζί, δεν θα ήταν εύλογο να περιοριστούμε στις επιμέρους επιδράσεις των τελευταίων.

Δεν θα επεκταθούμε σε περαιτέρω ανάλυση των δεικτών και των συσχετίσεων τους, καθώς έχουμε ήδη εξαντλήσει το θέμα όταν αναλύσαμε τις ίδιες σχεδόν μεταβλητές στο probit.

5.2.3. Στατιστική σημαντικότητα μοντέλου- Τεστ:

Αρχικό δείγμα (1yr prior)

Πίνακας στατιστικών μέτρων μοντέλου			
Mean dependent var	0.5	S.D. dependent var	0.503003
S.E. of regression	0.396336	McFadden R-squared	0.366754
Log likelihood	-36.87032		
LR statistic (5 df)	42.70808		
Probability (LR stat)	4.23E-08		

Παρατηρούμε ότι και εδώ ο μέσος και η τυπική απόκλιση του y είναι 0.5, τιμές που πάλι δικαιολογούνται από τις τιμές 0, 1 που έχουμε δώσει στο y .

Το τυπικό σφάλμα της παλινδρόμησης εμφανίζεται κι εδώ μειωμένο και κυμαίνεται στα επίπεδα του αντίστοιχου στο probit.

Το ίδιο συμβαίνει και με τη μέγιστη τιμή της συνάρτησης πιθανοφάνειας, κυμαίνεται και αυτή στα ίδια επίπεδα με του probit.

Η τιμή της πιθανότητας του LR statistic εμφανίζεται ελάχιστα μεγαλύτερη από αυτή του probit, πλησιάζοντας και αυτή όμως το μηδέν. Από τη στιγμή λοιπόν που είναι κατά πολύ μικρότερη του 0.05, μπορούμε να θεωρήσουμε ότι όλες οι εκτιμημένες παράμετροι είναι διάφορες του μηδενός και άρα το μοντέλο μας είναι στατιστικά σημαντικό.

Τέλος, το R^2 έχει την ίδια τιμή με αυτό του probit, 0.36.

5.2.4. Ακρίβεια πρόβλεψης μοντέλου:

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Logit			
Date: 05/11/02 Time: 20:27			
Sample (adjusted): 1 84			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P			
(Dep=1)≤C	33	10	43
P (Dep=1)>C	9	32	41
Total	42	42	84
Correct	33	32	65
% Correct	78.57	76.19	77.38
% Incorrect	21.43	23.81	22.62

Με τη βοήθεια του παραπάνω πίνακα διαπιστώνουμε ότι η ακρίβεια πρόβλεψης του μοντέλου είναι 77.38%, ακριβώς ίδια με αυτή του probit. Ίδια ακριβώς προκύπτουν και τα επιμέρους αποτελέσματα, οι επιμέρους δηλαδή ακρίβειες. Και εδώ υπερέχει η πρόβλεψη στις καθυστερημένες επιχειρήσεις, δίνοντας 33 στις 42 παρατηρήσεις σωστή πρόβλεψη και μόνο 9 λάθος. Ελάχιστα πιο κάτω όμως βρίσκεται και η ακρίβεια στις υγιείς, αφού προβλέπονται σωστά 32 υγιείς επιχειρήσεις ενώ λάθος μόνο 10. Οφείλουμε όμως πάλι να επισημάνουμε ότι η ακρίβεια κατάταξης εδώ αναμενόταν αρκετά ικανοποιητική, καθώς ο έλεγχος για τη διαπίστωσή της έγινε στο αρχικό δείγμα από το οποίο προήλθε το μοντέλο. Παρακάτω μεταβάλλουμε το cutoff score και παρατηρούμε την εξέλιξη της ακρίβειας:

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Logit			
Date: 05/11/02 Time: 20:27			
Sample (adjusted): 1 84			
Included observations: 84 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	29	6	35
P (Dep=1)>C	13	36	49
Total	42	42	84
Correct	29	36	65
% Correct	69.05	85.71	77.38
% Incorrect	30.95	14.29	22.62

Expectation prediction table			
Dependent Variable: Y			
Method: ML - Binary Logit			
Date: 05/11/02 Time: 20:27			
Sample (adjusted): 1 84			
Included observations: 84 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P (Dep=1)<=C	37	13	50
P (Dep=1)>C	5	29	34
Total	42	42	84
Correct	37	29	66
% Correct	88.1	69.05	78.57
% Incorrect	11.9	30.95	21.43

Παρατηρούμε ότι μία μείωση του cutoff point δεν μεταβάλλει την ακρίβεια του μοντέλου, απλά αυξάνει κατά πολύ την ακρίβεια πρόβλεψης στις υγιείς. Αντίθετα, η ακρίβεια πρόβλεψης στις καθυστερημένες μειώνεται κατά πολύ. Η αύξηση της ακρίβειας πρόβλεψης στις υγιείς υποδηλώνει ότι πολλές εταιρίες που παρουσίαζαν σκορ 0.4 έως 0.5 ήταν παρ' ολ' αυτά υγιείς. Το αντίτιμο όμως ενός τέτοιου cutoff point είναι ότι και κάποιες προβληματικές που εμφάνιζαν σκορ από 0.4 έως 0.5 λαμβάνονται πλέον ως υγιείς και κατατάσσονται λάθος.

Η αύξηση από την άλλη του cutoff score επιδρά θετικά στη συνολική ακρίβεια του μοντέλου αυξάνοντάς την. Αυξάνεται επίσης η ακρίβεια πρόβλεψης στις προβληματικές, κάτι που μας ενδιαφέρει πολύ, αλλά μειώνεται η αντίστοιχη των υγιών. Και στο logit λοιπόν το τεστ που γίνεται στο αρχικό δείγμα μας δίνει ως καλύτερο cutoff point το 0.6. Ας μην ξεχνάμε όμως, ότι ενώ συνέβη το ίδιο και με το probit, τα υπόλοιπα τεστ του τελευταίου ανέδειξαν ως καλύτερο σκορ το 0.4. Ενδιαφέρον έχει να δούμε αν θα συμβεί το ίδιο και εδώ.

5.2.5. Goodness of fit test:

Το τεστ αυτό συγκρίνει τις αναμενόμενες τιμές με τις πραγματικές ανά γκρουπ. Αν οι διαφορές που θα προκύψουν είναι μεγάλες, απορρίπτουμε το μοντέλο ως μη αντιπροσωπευτικό των αληθινών δεδομένων. Το Hosmer-Lemeshow τεστ που διεξήγαμε μας έδωσε μια χαμηλή τιμή της τάξεως του H-L Statistic: 6.0305 (λίγο μεγαλύτερη από του probit), που δεν μας επιτρέπει να απορρίψουμε το μοντέλο μας.

Στη συνέχεια παραθέτουμε μία σειρά από τεστ για την επιβεβαίωση της ακρίβειας του μοντέλου μας, τα ίδια ακριβώς που παραθέσαμε και στο probit :

Τεστ 1: Αρχικό δείγμα 2yrs prior

Διεξήγαμε το ίδιο τεστ με το probit: χρησιμοποιήσαμε ως βάση τα δεδομένα του αρχικού πάλι δείγματος αλλά για 2 έτη πριν την καθυστέρηση και εξετάσαμε την ακρίβεια πρόβλεψης του μοντέλου. Εισάγαμε δηλαδή ως δεδομένα τους 5 δείκτες των 54 πια εταιριών, τις παραμέτρους του μοντέλου μας που είναι πια γνωστές και επιλύσαμε ως προς το άγνωστο πλέον y . Στη συνέχεια, συγκρίναμε τα εκτιμημένα αυτά y με τα πραγματικά και διαπιστώσαμε το ποσοστό ακρίβειας του μοντέλου μας σε αυτό το δείγμα. Παρακάτω παρατίθεται ο πίνακας των αποτελεσμάτων με διαφορετικό κάθε φορά cutoff score:

test1			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 54			
Included observations: 54 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)≤C	23	6	29
P(Dep=1)>C	4	21	25
Total	27	27	54
Correct	23	21	44
% Correct	85.19	77.78	81.48
% Incorrect	14.81	22.22	18.52

test1	
Dependent Variable: Y	
Method: ML - Binary Logit	

Sample(adjusted): 1 54
 Included observations: 54 after adjusting endpoints
 Prediction Evaluation (success cutoff C = **0.4**)

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)<=C	23	6	29
P(Dep=1)>C	4	21	25
Total	27	27	54
Correct	23	21	44
% Correct	85.19	77.78	81.48
% Incorrect	14.81	22.22	18.52

test1
 Dependent Variable: Y
 Method: ML - Binary Logit
 Sample(adjusted): 1 54
 Included observations: 54 after adjusting endpoints
 Prediction Evaluation (success cutoff C = **0.6**)

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)<=C	23	6	29
P(Dep=1)>C	4	21	25
Total	27	27	54
Correct	23	21	44
% Correct	85.19	77.78	81.48
% Incorrect	14.81	22.22	18.52

Αυτό που παρατηρούμε είναι ότι σε αυτό το δείγμα, το μοντέλο παρουσιάζει ακόμη μεγαλύτερη ακρίβεια, της τάξεως του 81.48%, ακρίβεια αρκετά μεγαλύτερη από το 77.78% που παρουσίασε στο ίδιο τεστ το probit. Βέβαια, μία αρκετά καλή

ακρίβεια ήταν αναμενόμενη αφού το τεστ έγινε σε sub-sample του αρχικού δείγματος, αλλά δεν αναμενόταν και μεγαλύτερη από του αρχικού, καθώς το δείγμα αυτό αφορούσε παρατηρήσεις 2 ετών πριν την καθυστέρηση και όχι 1 όπως στο αρχικό. Η ακρίβεια στην πρόβλεψη των προβληματικών φτάνει μέχρι και το 85% περίπου, κάτι που μειώνει σημαντικά το Type I error, αλλά και η αντίστοιχη των υγιών κυμαίνεται στα επίσης υψηλά επίπεδα του 77.78%. Η μεταβολή του cutoff δεν ασκεί καμία απολύτως επίδραση στην ακρίβεια του μοντέλου.

Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior

Στη συνέχεια κάναμε το ίδιο ακριβώς τεστ στα στοιχεία όλων των εταιριών του αρχικού δείγματος και για το 1 και για τα 2 έτη πριν τη χρεωκοπία. Τα αποτελέσματα παρατίθενται πάλι για διαφορετικά cutoff scores:

test2			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 138			
Included observations: 138 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	60	18	78
P(Dep=1)>C	9	51	60
Total	69	69	138
Correct	60	51	111
% Correct	86.96	73.91	80.43
% Incorrect	13.04	26.09	19.57

Παρατηρούμε ότι η ακρίβεια του μοντέλου είναι επίσης αρκετά ικανοποιητική με πάντα υψηλότερη την ακρίβεια στην πρόβλεψη των προβληματικών. Και σε αυτό το τεστ το logit παρουσιάζει μεγαλύτερη ακρίβεια από το probit (77.54%).

Με τη μεταβολή του cutoff που παρουσιάζεται στους παρακάτω πίνακες παρατηρούμε ότι τόσο η μείωση όσο και η αύξησή του έχουν την ίδια επίδραση στη συνολική ακρίβεια του μοντέλου, δηλαδή τη μειώνουν. Αυτό είναι κάτι που το συναντάμε για πρώτη φορά. Η μείωση δηλαδή του cutoff έχει για πρώτη φορά ως αποτέλεσμα επίσης τη μείωση της συνολικής ακρίβειας. Η μείωση αυτή προέρχεται από τη μείωση στην ακρίβεια πρόβλεψης των προβληματικών, καθώς η αντίστοιχη των υγιών αυξάνεται αλλά δεν καταφέρνει να επηρεάσει τη συνολική. Επομένως για αυτό το τεστ καλύτερο cutoff score αποδεικνύεται το 0.5.

test2
 Dependent Variable: Y
 Method: ML - Binary Logit
 Sample (adjusted): 1 138
 Included observations: 138 after adjusting endpoints
 Prediction Evaluation (success cutoff C = **0.4**)

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)<=C	56	16	72
P(Dep=1)>C	13	53	66
Total	69	69	138
Correct	56	53	109
% Correct	81.16	76.81	78.99
% Incorrect	18.84	23.19	21.01

test2
 Dependent Variable: Y
 Method: ML - Binary Logit
 Sample(adjusted): 1 138
 Included observations: 138 after adjusting endpoints
 Prediction Evaluation (success cutoff C = **0.6**)

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)<=C	60	19	79
P(Dep=1)>C	9	50	59
Total	69	69	138
Correct	60	50	110
% Correct	86.96	72.46	79.71
% Incorrect	13.04	27.54	20.29

Τεστ 3: 2^ο δείγμα 2yrs prior

Αυτό όμως που πάλι, όπως και στο probit, έχει ιδιαίτερο ενδιαφέρον και θα μας διαφωτίσει περισσότερο για την αποτελεσματικότητα του μοντέλου μας είναι ο έλεγχος με δείγμα διαφορετικό από το αρχικό. Ο πρώτος έλεγχος έγινε στις παρατηρήσεις του δεύτερου δείγματος για 2 έτη πριν την καθυστέρηση. Οφείλουμε πάλι να επισημάνουμε ότι οι παρατηρήσεις αυτές δεν ήταν αρκετές ώστε το δείγμα μας να θεωρηθεί αντιπροσωπευτικό. Παρ' όλ' αυτά παραθέτουμε τα αποτελέσματα με διαφορετικά πάντα cutoff scores:

test3			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	9	3	12
P(Dep=1)>C	0	6	6
Total	9	9	18
Correct	9	6	15
% Correct	100.00	66.67	83.33
% Incorrect	0.00	33.33	16.67

test3			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	9	2	11
P(Dep=1)>C	0	7	7
Total	9	9	18
Correct	9	7	16
% Correct	100.00	77.78	88.89
% Incorrect	0.00	22.22	11.11

test3			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 18			
Included observations: 18 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	9	3	12
P(Dep=1)>C	0	6	6
Total	9	9	18
Correct	9	6	15
% Correct	100.00	66.67	83.33
% Incorrect	0.00	33.33	16.67

Παρατηρούμε ότι το συγκεκριμένο τεστ εμφανίζει το μοντέλο ως περισσότερο ακριβές από κάθε άλλο τεστ, με την ακρίβεια πρόβλεψης στις καθυστερημένες να φτάνει πάλι το 100% και τη συνολική ακρίβεια για πρώτη φορά στο 88.89% (με cutoff point 0.4), νούμερο που για πρώτη φορά εμφανίζεται σε οποιοδήποτε μοντέλο ή τεστ. Αλλά όπως είπαμε και για το probit, το δείγμα αυτό είναι πολύ μικρό και απλώς έτυχε να μην γίνει καμία λάθος πρόβλεψη για τις προβληματικές. Η αυξομείωση του cutoff score έχει τα αναμενόμενα αποτελέσματα, μία μείωσή του αυξάνει την ακρίβεια, ενώ μια αύξησή του αφού δεν τη μειώνει την αφήνει απλώς αμετάβλητη. Τα συμπεράσματα από αυτό το τεστ δεν μπορούν να είναι ιδιαίτερα σημαντικά, λόγω της αδυναμίας του δείγματος. Το μόνο που θα μπορούσαμε ίσως πάλι να τονίσουμε είναι η υπεροχή των αποτελεσμάτων και σε αυτό το τεστ του logit έναντι του probit.

Τεστ 4: 2^ο δείγμα 1yr prior

Ο πιο ενδιαφέρων και αντιπροσωπευτικός ίσως έλεγχος είναι αυτός που έγινε στο δεύτερο αυτό δείγμα για 1 έτος πριν την καθυστέρηση. Το δείγμα αυτό περιείχε 38 παρατηρήσεις και μας έδωσε τα παρακάτω αποτελέσματα:

test4			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	17	15	32
P(Dep=1)>C	2	4	6
Total	19	19	38
Correct	17	4	21
% Correct	89.47	21.05	55.26
% Incorrect	10.53	78.95	44.74

test4			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	17	14	31
P(Dep=1)>C	2	5	7
Total	19	19	38
Correct	17	5	22
% Correct	89.47	26.32	57.89
% Incorrect	10.53	73.68	42.11

test4			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 38			
Included observations: 38 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	17	16	33
P(Dep=1)>C	2	3	5
Total	19	19	38
Correct	17	3	20
% Correct	89.47	15.79	52.63
% Incorrect	10.53	84.21	47.37

Παρατηρούμε ότι η ακρίβεια πλέον μειώνεται κατά πολύ. Από το 77 και το 80% πέφτει στο 55% πλέον. Και στο probit είχαμε παρατηρήσει μια μείωση αλλά πολύ μικρότερη (από 77 σε 76% περίπου). Παρόλο που το logit σε όλα τα προηγούμενα τεστ εμφανιζόταν καλύτερο από άποψη ακρίβειας από το probit, σε αυτό το τεστ ξαφνικά η ακρίβειά του μειώνεται θεαματικά σε σχέση τόσο με το αντίστοιχο τεστ του probit όσο και με τα υπόλοιπα τεστ του ίδιου του logit. Δύο εκδοχές υπάρχουν ως ερμηνεία: η πρώτη είναι ότι το δείγμα αυτό ως διαφορετικό από το αρχικό δείγμα από το οποίο προέκυψε το μοντέλο είναι πιο αντιπροσωπευτικό για τη διεξαγωγή ελέγχων και επομένως τα αποτελέσματα που μας δίνει θα πρέπει να τα λάβουμε σοβαρά υπόψη. Η δεύτερη από την άλλη, λαμβάνει υπόψη ότι το συγκεκριμένο δείγμα πέρα από το ότι είναι διαφορετικό από το αρχικό, δεν τελεί όλα τα κριτήρια ενός αντιπροσωπευτικού δείγματος, καθώς η αντιστοίχιση από την οποία προέκυψε έγινε μόνο βάσει μεγέθους εταιριών και ετών. Μπορεί δηλαδή αυτό το δείγμα να περιέχει κάλλιστα ως αντιστοιχισμένα ζεύγη εταιρίες διαφορετικού κλάδου και αντικειμένου εργασιών και επομένως μη συγκρίσιμες. Η αδυναμία επομένως αυτή του δείγματος μπορεί να αποτελεί μια πολύ καλή αιτιολογία για την εμφάνιση των παραπάνω αποτελεσμάτων. Ο αντίλογος βέβαια σε αυτό το επιχείρημα είναι τα αποτελέσματα του αντίστοιχου τεστ στο probit. Αν το δείγμα μας πραγματικά δεν ήταν κατάλληλο για τη διεξαγωγή ελέγχου, θα έπρεπε κατά πάσα πιθανότητα να δείξει πολύ μικρότερη ακρίβεια και για το probit, κάτι που όμως δεν συμβαίνει.

Είμαστε λοιπόν αντιμέτωποι με δύο αντιφατικά αποτελέσματα για το logit, τα 3 προηγούμενα τεστ που δίνουν μεγάλη ακρίβεια για το μοντέλο και το 4^ο αυτό τεστ που δίνει πολύ μικρή. Πριν καταλήξουμε σε ένα συμπέρασμα, ας δούμε και τα αποτελέσματα του 5^{ου} τεστ:

Τεστ 5: Αρχικό δείγμα (1yr prior)- ΕΠΑΝΕΚΤΙΜΗΣΗ ΠΑΡΑΜΕΤΡΩΝ

Οι νέες επανεκτιμημένες παράμετροι του logit είναι οι εξής:

$$Y = - 5.7012 \cdot X_2 + 10.568 \cdot X_6 + 2.606 \cdot X_{10} + 2.156 \cdot X_{11} + 2.908 \cdot X_{13}$$

Τα αποτελέσματα από τον έλεγχο με βάση τις νέες παραμέτρους και το υπόλοιπο του αρχικού δείγματος είναι τα παρακάτω:

test5			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.5)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	20	11	31
P(Dep=1)>C	1	10	11
Total	21	21	42
Correct	20	10	30
% Correct	95.24	47.62	71.43
% Incorrect	4.76	52.38	28.57

test5			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.4)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	19	11	30
P(Dep=1)>C	2	10	12
Total	21	21	42
Correct	19	10	29
% Correct	90.48	47.62	69.05
% Incorrect	9.52	52.38	30.95

test5			
Dependent Variable: Y			
Method: ML - Binary Logit			
Sample(adjusted): 1 42			
Included observations: 42 after adjusting endpoints			
Prediction Evaluation (success cutoff C = 0.6)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	19	11	30
P(Dep=1)>C	2	10	12
Total	21	21	42
Correct	19	10	29
% Correct	90.48	47.62	69.05
% Incorrect	9.52	52.38	30.95

Παρατηρούμε ότι η ακρίβεια πάλι μειώνεται αλλά όχι τόσο πολύ όσο στο προηγούμενο τεστ ή όσο στο ίδιο τεστ στο probit (από 77.38% σε 66.67%). Εδώ έχουμε μια μείωση από 77.38% σε 71.43% . (Λαμβάνουμε πάντα ως μέτρο σύγκρισης την ακρίβεια που προέκυψε από το τεστ στο αρχικό δείγμα). Η αυξομείωση του cutoff point έχει την ίδια ακριβώς επίδραση στη συνολική ακρίβεια του μοντέλου, τη μείωσή της. Πάντα όμως υπερισχύει η ακρίβεια στην πρόβλεψη των καθυστερημένων. Η μείωση αυτή όμως της συνολικής ακρίβειας του logit στα δύο τελευταία τεστ οφείλει να μας προβληματίσει για την ορθότητα και καταλληλότητα του μοντέλου. Η αντίφαση αυτή μεταξύ των αποτελεσμάτων των πρώτων τεστ και των αποτελεσμάτων των τελευταίων δημιουργεί αμφιβολίες και υποψίες για την πραγματική ακρίβεια του μοντέλου. Ίσως η μεγάλη ακρίβεια που παρουσίασε το μοντέλο στα αρχικά τεστ να είναι απλά φαινομενική και παραπλανητική και ίσως αυτή που θα έπρεπε να λάβουμε πραγματικά υπόψη μας, είναι αυτή που προκύπτει από δείγματα άλλα εκτός του αρχικού.

Μετά από όλα αυτά τα τεστ νομίζουμε ότι μπορούμε πλέον να καταλήξουμε σε κάποια συμπεράσματα τόσο για το probit όσο και για το logit:

5.2.6.Γενικά συμπεράσματα για probit και logit:

Και τα δύο μοντέλα παρουσιάζουν μία μέση ακρίβεια της τάξεως του 75.6% το logit και 77.39% για το probit τόσο για το ένα όσο και για τα δύο χρόνια πριν την καθυστέρηση. Το logit αποδεικνύεται σε όλα σχεδόν τα τεστ εκτός του 4^{ου} καλύτερο από το probit και αυτό ακριβώς το τεστ είναι που κατεβάζει τον μέσο όρο ακρίβειας του logit. Και τα δύο μοντέλα παρουσιάζουν σε ένα από όλα τα τεστ προβληματική συμπεριφορά, δηλαδή μειωμένη προβλεψιμότητα. Το probit παρουσιάζει πρόβλημα στο τεστ 5 της επανεκτίμησης των παραμέτρων και το logit

στο τεστ 4 της χρησιμοποίησης διαφορετικού δείγματος. Το πιο ανησυχητικό από τα δύο είναι το πρόβλημα του logit, καθώς το συγκεκριμένο τεστ του probit μπορούμε κάλλιστα να θεωρήσουμε ότι έγινε με ένα ανεπαρκές σε μέγεθος δείγμα. Τις ερμηνείες μας και τους προβληματισμούς μας για το συγκεκριμένο πρόβλημα του logit τις παραθέσαμε παραπάνω. Αυτό που μένει να πούμε είναι ότι από τη στιγμή που όλα τα υπόλοιπα τεστ έδειξαν μια πολύ καλή ακρίβεια για το μοντέλο εκτός του συγκεκριμένου, οφείλουμε να θεωρήσουμε το μοντέλο ικανοποιητικό αλλά με επιφυλάξεις. Το probit εμφανίζεται ελαφρώς λιγότερο ικανοποιητικό από το logit στο σύνολο των τεστ, αλλά με μικρότερες διαφορές ανάμεσα σε αυτά. Ίσως λοιπόν θα έπρεπε να θεωρήσουμε το probit ως πιο ασφαλές μοντέλο, έστω και με ελαφρά μικρότερη ακρίβεια, σε σχέση με το φαινομενικά καλύτερο logit, που εξαιτίας ακριβώς αυτής της υπεροχής του θα μπορούσε να θεωρηθεί ύποπτο.

Ως σημαντικό πλεονέκτημα και των δύο μοντέλων θεωρούμε την έμφαση στην ακρίβεια πρόβλεψης των προβληματικών επιχειρήσεων, τη μείωση δηλαδή του Σφάλματος Τύπου I, αφού μας είναι πιο χρήσιμο να προβλέπουμε σωστά τις προβληματικές εταιρίες, καθώς αυτές θα μας προκαλέσουν ζημιά, από τις υγιείς.

Το cutoff score και για τα δύο μοντέλα αποδεικνύεται να είναι το 0.4 αφού αυτό μας δίνει μεγαλύτερη ακρίβεια στην πλειοψηφία των τεστ.

Γενικά, θα μπορούσαμε να πούμε ότι τα δύο αυτά μοντέλα προβλέπουν με ακρίβεια 75% περίπου την κατάσταση μιας επιχείρησης μέχρι και 2 χρόνια πριν την καθυστέρηση. Για ένα πιο σίγουρο όμως αποτέλεσμα θα έπρεπε να γίνουν πολλά τεστ σε πολλά διαφορετικά δείγματα, κάτι που εμείς λόγω έλλειψης στοιχείων δεν μπορέσαμε να κάνουμε.

Οφείλουμε τέλος να επισημάνουμε ότι όποια αποτελέσματα και αν βγάλαμε για τα δύο αυτά μοντέλα, για τα οποία πρέπει να ομολογήσουμε ότι τα θεωρούμε ικανοποιητικά δεδομένου των στοιχείων που διαθέταμε, δεν μπορούν να αποτελέσουν παρά μόνο ένα κριτήριο στο συμπέρασμα που θα βγάλει κάποιος που μελετά το μέλλον μιας επιχείρησης. Σε καμία περίπτωση δηλαδή τα μοντέλα αυτά, όσο μεγάλη ακρίβεια κι αν παρουσιάζουν, δεν θα έπρεπε να χρησιμοποιηθούν ως μοναδικά κριτήρια αξιολόγησης μίας ή πολλών εταιριών.

5.3.MDA

5.3.1.Υποθέσεις μοντέλου

Η discriminant analysis έχει το χαρακτηριστικό να προβλέπει διαχωρίζοντας τις παρατηρήσεις ανάμεσα σε δύο γκρουπ. Η διαδικασία έγκειται στη δημιουργία μιας discriminant function με βάση γραμμικούς συνδυασμούς των ανεξάρτητων μεταβλητών που παρέχουν τον καλύτερο διαχωρισμό ανάμεσα στα δύο γκρουπ. Η συνάρτηση δημιουργείται και εδώ όπως και στα δύο προηγούμενα μοντέλα, από το αρχικό δείγμα των 84 παρατηρήσεων, όπου η κατανομή στα γκρουπ είναι γνωστή. Στη συνέχεια, διεξάγουμε μια σειρά από τεστ για την διαπίστωση της ακρίβειας του μοντέλου, τα οποία στηρίζονται και πάλι στην λογική της εφαρμογής της συνάρτησης στο ίδιο και σε διαφορετικό δείγμα με την κατάταξη όμως του δείγματος (y) να εισάγεται ως ο ζητούμενος όρος της εξίσωσης και όλα τα υπόλοιπα να δίνονται ως δεδομένα.

Η discriminant analysis στηρίζεται σε μια σειρά από παραδοχές τις οποίες οφείλουμε να αναφέρουμε:

- Οι ανεξάρτητες μεταβλητές θεωρούμε ότι έχουν πολυμεταβλητή κανονική κατανομή.
- Οι πίνακες συνδιακυμάνσεων του κάθε γκρουπ ξεχωριστά (within-group variance-covariance matrices) θεωρούμε ότι είναι ίσοι.
- Η κατανομή σε κάθε γκρουπ θεωρείται αμοιβαίως αποκλειόμενη, δηλαδή καμία περίπτωση δεν μπορεί να ανήκει σε παραπάνω από δύο γκρουπ. Επίσης όλες οι περιπτώσεις πρέπει να είναι μέλη κάποιου από τα δύο γκρουπ.
- Για τη δημιουργία του μοντέλου θεωρήσαμε ότι οι prior probabilities είναι ίσες. Αυτό σημαίνει ότι η πιθανότητα να συναντήσουμε στον πληθυσμό μία υγιή επιχείρηση είναι ίση με την πιθανότητα να συναντήσουμε μία προβληματική. Γνωρίζουμε βέβαια ότι κάτι τέτοιο δεν ισχύει στην πράξη, καθώς οι υγιείς επιχειρήσεις μέσα στον πληθυσμό είναι πολύ περισσότερες από τις προβληματικές. Γνωρίζουμε επίσης ότι αυτό μπορεί να συνεπάγεται ένα potential bias. Παρ' όλ' αυτά, σύμφωνα και με όλες τις προηγούμενες μελέτες, τις λάβαμε ως ίσες γιατί θεωρήσαμε ότι το bias αφορά τη συχνότητα με την οποία θα συναντήσουμε σε έναν ολόκληρο πληθυσμό μία προβληματική επιχείρηση. Αντικείμενο όμως της συγκεκριμένης μελέτης είναι η εισαγωγή των στοιχείων μιας επιχείρησης από ένα δείγμα και η σωστή πρόβλεψη της κατάστασής της. Δεν ασχολούμαστε με άλλα λόγια με τη συνολική συχνότητα εμφάνισης μιας προβληματικής κατάστασης σε έναν ολόκληρο πληθυσμό, αλλά με την εξέταση μιας μεμονωμένης κάθε φορά περίπτωσης και τη σωστή πρόβλεψή της.

5.3.2. Στατιστική σημαντικότητα μεταβλητών:

Παρακάτω παρατίθενται τα αποτελέσματα που προκύπτουν από το πρόγραμμα SPSS κατά τη δημιουργία του μοντέλου:

ΠΙΝΑΚΑΣ 1

Analysis Case Processing Summary

Unweighted Cases		N	Percent
Valid		84	100.0
Excluded	Missing or out-of-range group codes	0	.0
	At least one missing discriminating variable	0	.0
	Both missing or out-of-range group codes and at least one missing discriminating variable	0	.0
	Total	0	.0
Total		84	100.0

Ο ΠΙΝΑΚΑΣ 1 δείχνει το σύνολο των παρατηρήσεων που συμπεριλήφθησαν στη διαδικασία δημιουργίας του μοντέλου. Βλέπουμε ότι χρησιμοποιήθηκαν όλες οι παρατηρήσεις που δόθηκαν, χωρίς καμία να μείνει εκτός.

Ο ΠΙΝΑΚΑΣ 2 παρουσιάζει τους μέσους και τις τυπικές αποκλίσεις των 13 επιμέρους μεταβλητών. Παρατηρούμε ότι μας δίνονται πρώτα οι μέσοι και οι τυπικές αποκλίσεις για τις μεταβλητές του γκρουπ των προβληματικών εταιριών, κατόπιν για τις μεταβλητές του γκρουπ των υγιών και τέλος για το σύνολο τους.

Ενδιαφέρον έχει να παρατηρήσουμε τη συμπεριφορά, αν όχι όλων, κάποιων τουλάχιστον μεταβλητών, καθώς μετακινούμαστε από το ένα γκρουπ στο άλλο.

Η x_2 , που στα δύο προηγούμενα μοντέλα παρουσίαζε όπως είδαμε αντίθετη από την αναμενόμενη συμπεριφορά, λόγω της συσχέτισής της με τις υπόλοιπες, εδώ παρουσιάζει την αναμενόμενη λογική συμπεριφορά. Καθώς κινούμαστε από το γκρουπ των προβληματικών στο γκρουπ των υγιών αυξάνεται. Είναι λογικό, καθώς μία αύξηση των ρευστοποιούμενων στοιχείων του Ενεργητικού της επιχείρησης ως προς το Συνολικό Ενεργητικό της, βελτιώνει την κατάστασή της. Βέβαια και εδώ, παρατηρούμε ότι η μεταβολή της από το ένα γκρουπ στο άλλο είναι ελάχιστη, κάτι που δείχνει μια τάση της μεταβλητής προς την αντίθετη συμπεριφορά που εμφανίζει στα άλλα δύο μοντέλα. Επομένως και η λογιστική και η στατιστική εξήγηση που δώσαμε νωρίτερα στα μοντέλα logit και probit για την αντίθετη συμπεριφορά του δείκτη δεν καταρρίπτεται τελείως από αυτό το μοντέλο, ούτε και επιβεβαιώνεται όμως.

Αντίθετη σίγουρα συμπεριφορά παρουσιάζει η x_1 , η οποία εμφανίζεται μικρότερη στις υγιείς εταιρίες από ότι στις προβληματικές. Μία αύξηση του Κυκλοφορούντος Ενεργητικού της επιχείρησης ως προς τις Βραχυπρόθεσμες υποχρεώσεις της, δηλαδή μια αύξηση του δείκτη θα έπρεπε να σημαίνει βελτίωση της κατάστασης της επιχείρησης και επομένως οι μεγαλύτερες τιμές του δείκτη θα έπρεπε να ανήκουν στις υγιείς επιχειρήσεις και όχι στις προβληματικές. Μία εξήγηση που θα μπορούσε να δοθεί για αυτήν τη συμπεριφορά είναι η ίδια λογιστική και στατιστική

εξήγηση που δόθηκε στα probit, logit για τον x_2 , καθώς πρόκειται για παρόμοιο δείκτη.

Αντίθετη επίσης συμπεριφορά παρουσιάζει και η x_4 (Παθητικό / Καθαρή Θέση), η οποία θα έπρεπε να αυξάνεται στις προβληματικές και να μειώνεται στις υγιείς. Η λογική υπαγορεύει ότι όσο πιο πολλά δάνεια έχει μια επιχείρηση (↑ Παθητικό) τόσο χειρότερη είναι η οικονομική της κατάσταση. Παρ' όλ' αυτά, μπορεί και μια υγιής επιχείρηση να έχει λάβει πολλά δάνεια, αλλά να τα χρησιμοποιεί ορθά και αποτελεσματικά για την αύξηση των κερδών της, να επιτυγχάνει με άλλα λόγια μόχλευση. Μία υγιής επιχείρηση δηλαδή μπορεί να είναι διαρκώς δανεισμένη σε μεγάλο βαθμό, αλλά να χρησιμοποιεί προς όφελος της τα ξένα αυτά κεφάλαια ώστε να επιτυγχάνει μεγάλα κέρδη τα οποία της δημιουργούν και μια εύρωστη οικονομική κατάσταση και τη συνεχή δυνατότητα να ξεπληρώνει τα χρέη της και να αναζητά καινούργιες χρηματοδοτήσεις. Επομένως, και μια υγιής επιχείρηση μπορεί να παρουσιάζει έναν αυξημένο τέτοιο δείκτη λόγω της αύξησης του Παθητικού της, δηλαδή του αριθμητή. Αλλά και με βάση τον παρονομαστή μπορούμε να εξηγήσουμε την αντίθετη συμπεριφορά του δείκτη. Θεωρητικά, μία προβληματική εταιρία έχει αυξημένο δείκτη όχι μόνο από το αυξημένο Παθητικό της αλλά και από τη μειωμένη Καθαρή της Θέση. Η τελευταία δηλαδή δεν επαρκεί για την κάλυψη του Παθητικού της. Αυτό όμως μπορεί να συμβαίνει και σε μια υγιή επιχείρηση αν αυτή είναι νεοϊδρυθείσα και δεν έχει προλάβει να αποθεματικά της και κατά συνέπεια την Καθαρή της Θέση. Αυτές είναι οι λογιστικές εξηγήσεις που μπορούμε να δώσουμε για τις παραπάνω μη αναμενόμενες συμπεριφορές των δύο δεικτών.

Από στατιστική άποψη, οι αντίθετες αυτές συμπεριφορές θα μπορούσαν και πάλι να ερμηνευτούν ως το αποτέλεσμα της επίδρασης των συσχετίσεων με τις υπόλοιπες.

Όλες οι υπόλοιπες μεταβλητές παρουσιάζουν την αναμενόμενη λογική συμπεριφορά, με εκείνες που περιέχουν κέρδη στον αριθμητή ($x_7, x_8, x_9, x_{10}, x_{12}, x_{13}$) να παρουσιάζονται όχι απλά μικρότερες αλλά αρνητικές στις προβληματικές και θετικές στις υγιείς.

ΠΙΝΑΚΑΣ2 Group Statistics					
Y		Mean	Std. Deviation	Valid N (listwise)	
				Unweighted	Weighted
0	X1	4.296233	20.63753809	42	42
	X2	0.519263	0.265965604	42	42
	X3	0.675077	0.24736292	42	42
	X4	3.595023	3.690125324	42	42
	X5	0.82751	0.32279805	42	42
	X6	-0.05742	0.204878173	42	42
	X7	-0.0094	0.100670886	42	42
	X8	0.020487	0.598707754	42	42
	X9	-0.39106	2.386291938	42	42
	X10	-0.4037	2.382319955	42	42
	X11	0.725897	0.542304151	42	42
	X12	-0.01416	0.094345486	42	42
	X13	-0.00838	0.547833525	42	42
1	X1	1.296865	0.847164876	42	42
	X2	0.552398	0.30168095	42	42
	X3	0.5308	0.28882291	42	42
	X4	4.25426	14.80779391	42	42
	X5	0.883622	0.229826864	42	42
	X6	0.070895	0.23452302	42	42
	X7	0.084586	0.122275891	42	42
	X8	0.403402	0.696126006	42	42
	X9	0.168395	0.608848386	42	42
	X10	0.112052	0.590057225	42	42
	X11	1.153305	0.696617581	42	42
	X12	0.06003	0.092085485	42	42
	X13	0.320861	0.595730562	42	42
Total	X1	2.796549	14.59516865	84	84
	X2	0.53583	0.283157098	84	84
	X3	0.602939	0.276946066	84	84
	X4	3.924641	10.73083927	84	84
	X5	0.855566	0.279928932	84	84
	X6	0.00674	0.228187019	84	84
	X7	0.037594	0.12094113	84	84
	X8	0.211944	0.673453894	84	84
	X9	-0.11133	1.753623851	84	84
	X10	-0.14582	1.744368602	84	84
	X11	0.939601	0.6566655	84	84
	X12	0.022936	0.099891127	84	84
	X13	0.156238	0.59244368	84	84

Οι τυπικές αποκλίσεις των περισσότερων μεταβλητών παρουσιάζονται αρκετά μικρές, εκτός των x_1 , x_4 , που είναι και οι δείκτες που παρουσίασαν την περίεργη συμπεριφορά. Το γεγονός αυτό σε συνδυασμό με την

αναπάντεχη συμπεριφορά τους αποτελεί ένδειξη ότι δεν θα έπρεπε να τους λάβουμε σοβαρά υπόψη μας. Το μοντέλο πράγματι δεν τους περιλαμβάνει.

Οι διαφορές ανάμεσα στους μέσους των δύο γκρουπ (between groups) και οι μικρές αποκλίσεις των τιμών των μεταβλητών από τη μέση τιμή του μέσα σε κάθε γκρουπ (within group) δίνουν αντίστοιχα τον διαχωρισμό που χρειαζόμαστε ανάμεσα στα 2 γκρουπ καθώς και την ομοιομορφία και ομοιογένεια του κάθε γκρουπ ξεχωριστά, ώστε να καθιστούν τις μεταβλητές αποτελεσματικές για την ένταξή τους στην Discriminant function, εκτός πάντα από τις x_1, x_4 . Παρατηρούμε ότι οι δείκτες με τους πιο διαφορετικούς μέσους ανάμεσα στα δύο γκρουπ και τις μικρότερες τυπικές αποκλίσεις μέσα σε κάθε γκρουπ είναι οι $x_3, x_6, x_7, x_{11}, x_{12}, x_{13}$. Οι περισσότεροι από αυτούς τους δείκτες θα θεωρηθούν κατάλληλοι και από το πρόγραμμα όπως θα δούμε παρακάτω για τη δοκιμαστική, αν όχι την οριστική, ένταξή τους στη συνάρτηση.

-Συσχέτιση μεταβλητών:

Ο ΠΙΝΑΚΑΣ 3 παρουσιάζει τις διακυμάνσεις, συνδιακυμάνσεις και συσχετίσεις ανάμεσα στους 13 δείκτες μέσα σε κάθε γκρουπ. Το επάνω τμήμα του πίνακα δείχνει τις συνδιακυμάνσεις μέσα στα γκρουπ (covariances within groups). Η διαγώνιος γραμμή αυτού του τμήματος εμφανίζει τις διακυμάνσεις καθεμιάς μεταβλητής. Το κάτω τμήμα του πίνακα δείχνει τις συσχετίσεις των μεταβλητών μέσα στα γκρουπ. Ο pooled within-groups πίνακας είναι τελείως διαφορετικός από τους συνηθισμένους πίνακες συνδιακυμάνσεων, όπου όλες οι περιπτώσεις αντιμετωπίζονται ως ξεχωριστό δείγμα. Οι pooled within-groups συνδιακυμάνσεις υπολογίζονται με βάση το μέσο όρο των ξεχωριστών πινάκων συνδιακυμάνσεων για κάθε γκρουπ. Οι pooled within-groups συσχετίσεις υπολογίζονται από τις διακυμάνσεις και τις συνδιακυμάνσεις.

Η Discriminant analysis ως μοντέλο, ορίζει ότι οι διακυμάνσεις των μεταβλητών μέσα στα γκρουπ (variances within groups) πρέπει να είναι όσο το δυνατόν μικρότερες ώστε να εξασφαλίζεται η ομοιογένεια του κάθε γκρουπ. Όσο πιο μικρή διακύμανση παρουσιάζουν δηλαδή οι μεταβλητές μέσα σε ένα γκρουπ τόσο πιο πολύ "ανήκουν" σε αυτό το γκρουπ. Οι μεταβλητές λοιπόν που παρουσιάζουν τις μικρότερες διακυμάνσεις είναι και οι υποψήφιες για την ένταξή τους στο μοντέλο. Τις μικρότερες διακυμάνσεις παρουσιάζουν με σειρά από τη μικρότερη προς τη μεγαλύτερη οι μεταβλητές: $x_{12}, x_7, x_6, x_3, x_5, x_2, x_{13}, x_{11}$.

Όσον αφορά τώρα τις συνδιακυμάνσεις και τις συσχετίσεις στο κάτω μέρος του πίνακα, παρατηρούμε ότι οι x_1, x_3, x_6 παρουσιάζουν αρνητική συσχέτιση με τις περισσότερες από τις υπόλοιπες, αποτέλεσμα διαφορετικό από αυτό που είχαμε δει στα probit και logit, όπου εκεί μόνο η x_2 παρουσίαζε αρνητική συσχέτιση. Η αρνητική συσχέτιση της x_1 ερμηνεύτηκε παραπάνω, όταν θελήσαμε να αναλύσουμε την αντίθετη συμπεριφορά της. Η αρνητική

συσχέτιση της x_3 δικαιολογεί και το αρνητικό της πρόσημο στην εξίσωση. Για τη συγκεκριμένη μεταβλητή, μπορούμε να δώσουμε μόνο μια στατιστική εξήγηση του αρνητικού της προσήμου, καθώς η συμπεριφορά της αν παρατηρήσουμε τον παραπάνω πίνακα των μέσων εμφανίζεται κανονική, αυξάνεται στις προβληματικές και μειώνεται στις υγιείς. Επομένως η μόνη αιτία στην οποία μπορούμε να αποδώσουμε το αρνητικό πρόσημο της x_3 είναι η αρνητική της συσχέτιση με τις περισσότερες μεταβλητές.

Όσο μεγαλύτερη συσχέτιση παρουσιάζει μια μεταβλητή με όλες τις υπόλοιπες τόσο καταλληλότερη είναι για να ενταχθεί στην τελική εξίσωση, καθώς θέλουμε η συσχέτιση των μεταβλητών μέσα σε κάθε γκρουπ να είναι μεγάλη, για να ανήκουν όλες στο γκρουπ. Παρατηρούμε ότι οι μεταβλητές $x_3, x_6, x_7, x_{11}, x_{13}$ που είναι οι υποψήφιοι για την ένταξη στο μοντέλο όπως θα δούμε παρακάτω, παρουσιάζουν μερικές από τις μεγαλύτερες συσχετίσεις μεταξύ τους. Υπάρχουν φυσικά και μερικές άλλες συσχετίσεις του ίδιου επιπέδου, αλλά οι μεταβλητές που τις παρουσιάζουν θα απορριφθούν από επόμενα κριτήρια.

ΠΙΝΑΚΑΣ3														
Pooled Within-Groups Matrices														
	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13	
Covariance	X1	213.3128	-0.5784	-0.7294	-6.36801	-1.42042	0.461508	0.007593	-0.07205	0.124607	0.124821	-1.2135	0.014784	-0.01298
	X2	-0.5784	0.080875	0.041095	0.915785	0.047836	0.015156	0.006275	0.013738	0.099458	0.099166	0.078299	0.004799	0.00706
	X3	-0.7294	0.041095	0.072304	1.033027	0.015633	-0.02921	-0.00041	0.079391	-0.10366	-0.10074	0.087116	-0.00095	0.06953
	X4	-6.36801	0.915785	1.033027	116.4439	0.480625	-0.39063	0.029137	2.692951	1.425326	1.508971	1.282801	0.049123	2.598081
	X5	-1.42042	0.047836	0.015633	0.480625	0.078509	-0.00958	0.007167	-0.0033	0.148807	0.148198	0.065644	0.005917	-0.0077
	X6	0.461508	0.015156	-0.02921	-0.39063	-0.00958	0.048488	-0.00102	-0.04469	-0.01303	-0.01593	-0.03947	-0.00151	-0.04043
	X7	0.007593	0.006275	-0.00041	0.029137	0.007167	-0.00102	0.012543	0.029889	0.114886	0.111097	0.02311	0.010168	0.020445
	X8	-0.07205	0.013738	0.079391	2.692951	-0.0033	-0.04469	0.029889	0.421521	-0.2577	-0.26582	0.11128	0.022131	0.366135
	X9	0.124607	0.099458	-0.10366	1.425326	0.148807	-0.01303	0.114886	-0.2577	3.032543	3.020023	0.123507	0.109881	-0.27625
	X10	0.124821	0.099166	-0.10074	1.508971	0.148198	-0.01593	0.111097	-0.26582	3.020023	3.011808	0.115657	0.107678	-0.28014
	X11	-1.2135	0.078299	0.087116	1.282801	0.065644	-0.03947	0.02311	0.11128	0.123507	0.115657	0.389685	0.016748	0.091539
	X12	0.014784	0.004799	-0.00095	0.049123	0.005917	-0.00151	0.010168	0.022131	0.109881	0.107678	0.016748	0.00869	0.015712
	X13	-0.01298	0.00706	0.06953	2.598081	-0.0077	-0.04043	0.020445	0.366135	-0.27625	-0.28014	0.091539	0.015712	0.327508
Correlation	X1	1	-0.13926	-0.18573	-0.04041	-0.3471	0.1435	0.004642	-0.0076	0.004899	0.004925	-0.1331	0.010858	-0.00155
	X2	-0.13926	1	0.537402	0.298421	0.600323	0.242027	0.197008	0.074406	0.20083	0.20093	0.441058	0.181029	0.043378
	X3	-0.18573	0.537402	1	0.356019	0.207492	-0.49334	-0.01364	0.454762	-0.22138	-0.21587	0.51899	-0.03794	0.451835
	X4	-0.04041	0.298421	0.356019	1	0.15896	-0.16439	0.024109	0.38438	0.075849	0.080577	0.190434	0.048833	0.42071
	X5	-0.3471	0.600323	0.207492	0.15896	1	-0.15535	0.228385	-0.01815	0.304971	0.304767	0.375298	0.226532	-0.04801
	X6	0.1435	0.242027	-0.49334	-0.16439	-0.15535	1	-0.04148	-0.3126	-0.03398	-0.0417	-0.28711	-0.07358	-0.32087
	X7	0.004642	0.197008	-0.01364	0.024109	0.228385	-0.04148	1	0.411054	0.589064	0.571595	0.330557	0.973947	0.318986
	X8	-0.0076	0.074406	0.454762	0.38438	-0.01815	-0.3126	0.411054	1	-0.22793	-0.23592	0.274567	0.36566	0.985417
	X9	0.004899	0.20083	-0.22138	0.075849	0.304971	-0.03398	0.589064	-0.22793	1	0.999294	0.113614	0.676863	-0.2772
	X10	0.004925	0.20093	-0.21587	0.080577	0.304767	-0.0417	0.571595	-0.23592	0.999294	1	0.106758	0.665569	-0.28206
	X11	-0.1331	0.441058	0.51899	0.190434	0.375298	-0.28711	0.330557	0.274567	0.113614	0.106758	1	0.287803	0.256235
	X12	0.010858	0.181029	-0.03794	0.048833	0.226532	-0.07358	0.973947	0.36566	0.676863	0.665569	0.287803	1	0.29451
	X13	-0.00155	0.043378	0.451835	0.42071	-0.04801	-0.32087	0.318986	0.985417	-0.2772	-0.28206	0.256235	0.29451	1

a The covariance matrix has 82 degrees of freedom.

Ο ΠΙΝΑΚΑΣ 4 είναι ο πίνακας συνδιακυμάνσεων του κάθε γκρουπ ξεχωριστά. Έχουμε ήδη αναφέρει ότι μία από τις υποθέσεις της Discriminant analysis είναι ότι οι πίνακες συνδιακυμάνσεων των δύο γκρουπ είναι ίσοι, ώστε η συνάρτηση να είναι γραμμική ως προς το x και όχι τετραγωνική. Ο παρακάτω πίνακας αποτελεί μια ένδειξη αυτού ακριβώς, αν οι πίνακες συνδιακυμάνσεων είναι ίσοι. Μας επιτρέπει να ψάξουμε για ίσες διακυμάνσεις και συνδιακυμάνσεις ανάμεσα στα γκρουπ. Το πλήθος βέβαια των μεταβλητών και των αριθμών δυσχεραίνουν τη διαδικασία διαπίστωσης αυτής της ισότητας. Οι διακυμάνσεις παρατίθενται στην κύρια διαγώνιο και με μια καλή ματιά, παρατηρούμε ότι οι διακυμάνσεις των $x_2, x_3, x_6, x_7, x_8, x_{11}, x_{12}, x_{13}$ παρουσιάζουν ελάχιστη απόκλιση ανάμεσα στα δύο γκρουπ, κάτι που αποτελεί σοβαρή ένδειξη ότι οι πίνακες συνδιακυμάνσεων των συγκεκριμένων μεταβλητών είναι περίπου ίσοι. Για τις ίδιες περίπου μεταβλητές παρατηρούμε ότι οι περισσότερες συνδιακυμάνσεις τους με τις υπόλοιπες μεταβλητές παρουσιάζουν ομοιότητα ανάμεσα στα δύο γκρουπ. Είναι αρκετά δύσκολο να διαπιστώσουμε με ακρίβεια την ισότητα των πινάκων συνδιακυμάνσεων από αυτόν τον πίνακα, αλλά δυστυχώς δεν υπάρχει κάποιο απλό τεστ για να ελέγξουμε τη συγκεκριμένη υπόθεση.

-Stepwise statistics:

Ο ΠΙΝΑΚΑΣ 5 παρουσιάζει τη διαδικασία και τα βασικά κριτήρια με τα οποία έγινε η τελική επιλογή των μεταβλητών. Οι προηγούμενοι πίνακες μας βοήθησαν να δούμε ποιες ήταν οι καταλληλότερες μεταβλητές για το μοντέλο με βάση τους μέσους, τις τυπικές αποκλίσεις, τις διακυμάνσεις, τις συνδιακυμάνσεις και τις συσχετίσεις τους. Με βάση όλα αυτά τα στατιστικά κριτήρια, το πρόγραμμα επέλεξε 5 μεταβλητές από τις οποίες τελικά κατέληξε σε 3 μετά από μια διαδικασία διαδοχικών βημάτων που του ζητήσαμε να κάνει (stepwise statistics).

Η πρώτη στήλη του πίνακα περιέχει τον αριθμό κάθε βήματος. Βλέπουμε ότι η διαδικασία έγινε σε 7 βήματα.

Η δεύτερη στήλη παρουσιάζει τις μεταβλητές που εισέρχονταν στο μοντέλο σε κάθε βήμα.

Η τρίτη στήλη παρουσιάζει τις μεταβλητές που χρειάστηκε να εξέλθουν από το μοντέλο.

Παρατηρούμε ότι στο 1^ο βήμα το πρόγραμμα δοκιμάζει μόνο τη μεταβλητή x_7 , στο 2^ο βήμα εισάγει και τη x_6 , στο 3^ο και τη x_{11} , στο 4^ο και τη x_3 , στο 5^ο βήμα αποκλείει τη x_7 , στο 6^ο αποκλείει και τη x_6 και τέλος στο 7^ο εισάγει και τη x_{13} . Το τελικό αποτέλεσμα είναι ένα μοντέλο που περιέχει τελικά τρεις μόνο μεταβλητές, τις x_3, x_{11}, x_{13} . Ας εξηγήσουμε τώρα τα κριτήρια με τα οποία έγιναν αυτές οι επιλογές.

Το βασικό κριτήριο είναι το Wilk's criterion ή αλλιώς το Wilk's lambda, που αποκαλείται και U statistic. Πρόκειται για έναν multivariate έλεγχο στατιστικής σημαντικότητας, που ελέγχει τις διαφορές ανάμεσα στα group centroids, δηλαδή το μέγεθος του διαχωρισμού ανάμεσα στα 2 γκρουπ (group discrimination). Πιο συγκεκριμένα, είναι το ποσοστό της συνολικής διακύμανσης των Discriminant scores που δεν εξηγείται από τις διαφορές ανάμεσα στα 2 γκρουπ. Επομένως, όσο μεγαλύτερο είναι τόσο μεγαλύτερο

ποσοστό της διακύμανσης δεν θα ερμηνεύεται, δεν θα οφείλεται στις διαφορές των γκρουπ, κάτι που δεν θέλουμε. Επιθυμούμε δηλαδή το μέτρο αυτό να είναι όσο το δυνατόν πιο μικρό. Η τιμή του κυμαίνεται μεταξύ 0 και 1, με τιμές κοντά στο 0 να συνεπάγονται ότι οι μέσοι των γκρουπ διαφέρουν και με τιμές κοντά στο 1 να δείχνουν το αντίθετο. Ένα παράλληλο κριτήριο είναι το F-ratio, το οποίο εκφράζει τη δεσμευμένη πιθανότητα ότι μια σχέση τόσο δυνατή όσο αυτή που παρατηρείται με τα συγκεκριμένα στοιχεία, θα ισχύει, αν η μηδενική υπόθεση αληθεύει. Αυτό που μας ενδιαφέρει κυρίως είναι η p-value της οποίας μια τιμή μικρότερη συνήθως του 0.05 σημαίνει ότι οι μεταβλητές μας είναι στατιστικά σημαντικές και οπότε τις δεχόμαστε.

Η μεταβλητή που μεγιστοποιεί το F-ratio ταυτόχρονα ελαχιστοποιεί το Wilk's lambda. Επομένως επιλέγουμε εκείνες τις μεταβλητές που εμφανίζουν τα μικρότερα Wilk's lambda και ταυτόχρονα τα μεγαλύτερα F-ratio.

Η τέταρτη λοιπόν στήλη μας δείχνει την τιμή του Wilk's lambda σε κάθε βήμα. Παρατηρούμε ότι τη μικρότερη τιμή τη λαμβάνουμε στο 7^ο τελικά βήμα.

Η πέμπτη στήλη παρουσιάζει τις αντίστοιχες τιμές του F-ratio. Παρατηρούμε ότι η μεγαλύτερη τιμή επιτυγχάνεται στο 6^ο βήμα. Παρ' όλ' αυτά, επιλέγουμε τελικά τις μεταβλητές του 7^{ου} βήματος, καθώς μόνο σε αυτό το βήμα επιτυγχάνουμε ταυτόχρονη ελαχιστοποίηση του lambda και έστω τη δεύτερη μεγαλύτερη τιμή του R.

Η έκτη τέλος στήλη παρουσιάζει τις τιμές της significance του F-ratio για τις οποίες οι μεταβλητές πρέπει να μείνουν ή να απορριφθούν. Οι υποσημειώσεις κάτω από τον πίνακα επισημαίνουν ότι όταν significance F-ratio < 0.05 η μεταβλητή πρέπει να εισαχθεί στο μοντέλο. Όταν significance F-ratio > 0.10 η μεταβλητή πρέπει να απορριφθεί. Παρατηρούμε ότι σε όλα τα βήματα η τιμή είναι μικρότερη του 0.05, αλλά τη μικρότερη και για αυτό το μέτρο την λαμβάνουμε στο 7^ο βήμα.

Ο ΠΙΝΑΚΑΣ 6 μας δείχνει τρία στατιστικά μέτρα για καθεμία από τις μεταβλητές που εισάγονται σε κάθε βήμα στο μοντέλο.

Στην πρώτη στήλη έχουμε και πάλι τα βήματα και τις μεταβλητές που περιλαμβάνει καθένα από αυτά.

Στη δεύτερη στήλη βλέπουμε ένα άλλο μέτρο στατιστικής σημαντικότητας των μεταβλητών που ονομάζεται tolerance. Είναι ένα μέτρο που ελέγχει τον βαθμό στον οποίο οι ανεξάρτητες μεταβλητές είναι γραμμικά συσχετισμένες μεταξύ τους (multicollinear). Δείχνει στην ουσία το ποσοστό της διακύμανσης μιας μεταβλητής που δεν ερμηνεύεται από τις υπόλοιπες ανεξάρτητες μεταβλητές της εξίσωσης. Μία μεταβλητή με πολύ χαμηλή tolerance συνεισφέρει λίγη πληροφορία για το μοντέλο και μπορεί να προκαλέσει υπολογιστικά προβλήματα. Υπολογίζεται ως $1-R^2$ για μια ανεξάρτητη μεταβλητή όταν προβλέπεται από τις υπόλοιπες ανεξάρτητες μεταβλητές που συμπεριλαμβάνονται ήδη στην ανάλυση.

Παρατηρούμε ότι στο 7^ο βήμα επιτυγχάνεται ένα ικανοποιητικό επίπεδο tolerance και για τις τρεις μεταβλητές. Και πάλι, στοχεύουμε στον καλύτερο δυνατό συνδυασμό πολλών κριτηρίων μαζί, με πρώτο πάντα κριτήριο το Wilk's κριτήριο. Σε αυτό λοιπόν το βήμα επιτυγχάνεται ο μικρότερος μέσος όρος του Wilk's lambda, ένα πολύ καλό ελάχιστο επίπεδο του significance of F και ένα επίσης πολύ καλό μέγιστο επίπεδο του tolerance.

Στο παράρτημα παρατίθεται και ο ΠΙΝΑΚΑΣ 2 που δείχνει τα ίδια στατιστικά μέτρα για τις μεταβλητές που δεν συμπεριλήφθησαν στο μοντέλο. Τα 7 βήματα έγιναν και για αυτές, απλά δεν συμπεριλήφθησαν τελικά στο μοντέλο γιατί τα παραπάνω στατιστικά μέτρα τους δεν ήταν ικανοποιητικά. Ενδιαφέρον έχει να δούμε πώς από το πρώτο βήμα έως το έβδομο στοχεύει και επιτυγχάνει το πρόγραμμα τη σταδιακή μείωση του Wilk's lambda.

ΠΙΝΑΚΑΣ4		Covariance Matrices												
Υ		X1	X2	X3	X4	X5	X6	X7	X8	X9	X10	X11	X12	X13
0	X1	425.908	-1.22815	-1.36657	-11.3132	-2.84959	0.777864	0.015424	-0.01753	0.144276	0.154363	-2.38563	0.030577	0.083768
	X2	-1.22815	0.070738	0.03037	0.58406	0.057601	0.00799	0.006254	-0.03693	0.182072	0.183	0.061776	0.006118	-0.03794
	X3	-1.36657	0.03037	0.061188	0.476436	0.018485	-0.02411	-0.00535	0.047772	-0.22072	-0.21986	0.067706	-0.00586	0.042867
	X4	-11.3132	0.58406	0.476436	13.61702	0.558959	-0.27869	0.096552	-0.62121	3.145365	3.152615	0.782348	0.089209	-0.69957
	X5	-2.84959	0.057601	0.018485	0.558959	0.104199	-0.0199	0.010912	-0.04273	0.328305	0.332274	0.080191	0.011002	-0.04237
	X6	0.777864	0.00799	-0.02411	-0.27869	-0.0199	0.041975	-0.00228	-0.0277	-0.03931	-0.04098	-0.04533	-0.00211	-0.02548
	X7	0.015424	0.006254	-0.00535	0.096552	0.010912	-0.00228	0.010135	0.008813	0.188323	0.186945	0.013137	0.009398	0.00363
	X8	-0.01753	-0.03693	0.047772	-0.62121	-0.04273	-0.0277	0.008813	0.358451	-0.59977	-0.60737	0.066025	0.004007	0.322301
	X9	0.144276	0.182072	-0.22072	3.145365	0.328305	-0.03931	0.188323	-0.59977	5.694389	5.683972	0.143677	0.18516	-0.61921
	X10	0.154363	0.183	-0.21986	3.152615	0.332274	-0.04098	0.186945	-0.60737	5.683972	5.675448	0.145355	0.184199	-0.62406
	X11	-2.38563	0.061776	0.067706	0.782348	0.080191	-0.04533	0.013137	0.066025	0.143677	0.145355	0.294094	0.012443	0.061275
	X12	0.030577	0.006118	-0.00586	0.089209	0.011002	-0.00211	0.009398	0.004007	0.18516	0.184199	0.012443	0.008901	0.000632
	X13	0.083768	-0.03794	0.042867	-0.69957	-0.04237	-0.02548	0.00363	0.322301	-0.61921	-0.62406	0.061275	0.000632	0.300122
1	X1	0.717688	0.071346	-0.09223	-1.42282	0.008739	0.145151	-0.00024	-0.12658	0.104938	0.09528	-0.04137	-0.00101	-0.10972
	X2	0.071346	0.091011	0.05182	1.247509	0.038071	0.022322	0.006295	0.064401	0.016843	0.015333	0.094823	0.00348	0.052056
	X3	-0.09223	0.05182	0.083419	1.589617	0.012781	-0.03431	0.004533	0.111011	0.013396	0.018389	0.106525	0.003962	0.096193
	X4	-1.42282	1.247509	1.589617	219.2708	0.402291	-0.50257	-0.03828	6.007112	-0.29471	-0.13467	1.783254	0.009037	5.89573
	X5	0.008739	0.038071	0.012781	0.402291	0.05282	0.000728	0.003421	0.036129	-0.03069	-0.03588	0.051097	0.000833	0.026975
	X6	0.145151	0.022322	-0.03431	-0.50257	0.000728	0.055001	0.000235	-0.06169	0.013254	0.009109	-0.0336	-0.00091	-0.05539
	X7	-0.00024	0.006295	0.004533	-0.03828	0.003421	0.000235	0.014951	0.050965	0.041449	0.035249	0.033084	0.010939	0.03726
	X8	-0.12658	0.064401	0.111011	6.007112	0.036129	-0.06169	0.050965	0.484591	0.084367	0.075742	0.156534	0.040255	0.409968
	X9	0.104938	0.016843	0.013396	-0.29471	-0.03069	0.013254	0.041449	0.084367	0.370696	0.356074	0.103336	0.034603	0.066697
	X10	0.09528	0.015333	0.018389	-0.13467	-0.03588	0.009109	0.035249	0.075742	0.356074	0.348168	0.085959	0.031157	0.06379
	X11	-0.04137	0.094823	0.106525	1.783254	0.051097	-0.0336	0.033084	0.156534	0.103336	0.085959	0.485276	0.021054	0.121803
	X12	-0.00101	0.00348	0.003962	0.009037	0.000833	-0.00091	0.010939	0.040255	0.034603	0.031157	0.021054	0.00848	0.030792
	X13	-0.10972	0.052056	0.096193	5.89573	0.026975	-0.05539	0.03726	0.409968	0.066697	0.06379	0.121803	0.030792	0.354895

ΠΙΝΑΚΑΣ5 Variables Entered/Removed					
Step	Entered	Removed	Wilks' Lambda Statistic	Exact F Statistic	Sig.F
1	X7		0.847	14.789	0.000237
2	X6		0.782	11.266	4.82E-05
3	X11		0.723	10.199	9.24E-06
4	X3		0.670	9.715	1.92E-06
5		X6	0.692	11.848	1.69E-06
6		X7	0.715	16.155	1.25E-06
7	X13		0.625	15.984	3.12E-08

At each step, the variable that minimizes the overall Wilks' Lambda is entered.

a Maximum number of steps is 26.

b Maximum significance of F to enter is .05.

c Minimum significance of F to remove is .10.

ΠΙΝΑΚΑΣ6 Variables in the Analysis				
Step		Tolerance	Sig. of F to Remove	Wilks' Lambda
1	X7	1	0.00024	
2	X7	0.998	0.00030	0.920
	X6	0.998	0.01135	0.847
3	X7	0.888	0.01868	0.775
	X6	0.914	0.00215	0.814
	X11	0.816	0.01252	0.782
4	X7	0.842	0.10416	0.693
	X6	0.754	0.11049	0.692
	X11	0.616	0.00084	0.772
	X3	0.571	0.01445	0.723
5	X7	0.844	0.11102	0.715
	X11	0.617	0.00072	0.799
	X3	0.692	0.00033	0.814
6	X11	0.731	0.00000	0.931
	X3	0.731	0.00002	0.893
7	X11	0.730	0.00003	0.776
	X3	0.622	0.00000	0.859
	X13	0.795	0.00110	0.715

Καταφέραμε λοιπόν με μία σειρά από στατιστικά κριτήρια όπως είναι οι μέσοι, οι διακυμάνσεις και συνδιακυμάνσεις, οι τυπικές αποκλίσεις και οι συσχετίσεις, το Wilk's lambda, το F-ratio, η significance του και η tolerance να καταλήξουμε σε μία linear Discriminant function με τρεις μεταβλητές: τις x_3 , x_{11} , x_{13} . Οι μεταβλητές αυτές θεωρήθηκαν με βάση τα παραπάνω

κριτήρια ως αυτές με τη μεγαλύτερη discriminating power και οι καταλληλότερες για την δημιουργία του μοντέλου.

Αυτό που μας ενδιαφέρει τώρα είναι να ελέγξουμε την ακρίβεια πρόβλεψης και την discriminating ability του μοντέλου μας, να ελέγξουμε δηλαδή κατά πόσον το μοντέλο που δημιουργήσαμε είναι στο σύνολό του ικανό και ακριβές.

5.3.3. Στατιστική σημαντικότητα μοντέλου- Τεστ :

Ένα μέτρο της στατιστικής σημαντικότητας του μοντέλου είναι η canonical correlation. Η canonical correlation δείχνει την ικανότητα της Discriminant function να διακρίνει ανάμεσα στα δύο γκρουπ (discriminating ability). Θα μπορούσαμε να ερμηνεύσουμε το τετράγωνό της ως το ποσοστό της διακύμανσης μέσα στην Discriminant function που ερμηνεύεται από τα γκρουπ. Μετράει δηλαδή τη σύνδεση μεταξύ των Discriminant scores και των groups, πόσο δηλαδή τα scores είναι πράγματι αντιπροσωπευτικά του κάθε γκρουπ. Τιμές κοντά στο 1 δείχνουν μια δυνατή συσχέτιση ανάμεσα στα Discriminant scores και τα groups. Η canonical correlation στο δικό μας μοντέλο είναι ίση με 0.612, μία τιμή όχι πάρα πολύ κοντά στο 1, αλλά τουλάχιστον ικανοποιητική.

Canonical Correlation
0.61219

Ο ΠΙΝΑΚΑΣ 7 παρουσιάζει τις παραμέτρους των μεταβλητών που επιλέχτηκαν για το μοντέλο σε standardized όρους, οι οποίοι μας βοηθούν να δούμε τη σχετική συνεισφορά της κάθε μεταβλητής στον συνολικό διαχωρισμό.

ΠΙΝΑΚΑΣ7	Standardized Canonical Discriminant Function Coefficients
	Function 1
X3	-1.080954723
X11	0.842336043
X13	0.648628249

Ο ΠΙΝΑΚΑΣ 8 παρουσιάζει τους μέσους του y για καθένα από τα δύο γκρουπ. Ο μέσος όρος του Discriminant score για το γκρουπ των προβληματικών είναι

-0.7649, ενώ ο αντίστοιχος για το γκρουπ των υγιών είναι 0.7649.

ΠΙΝΑΚΑΣ8	Functions at Group Centroids
	Function

Y	1
0	-0.764947385
1	0.764947385

Non-standardized canonical discriminant functions evaluated at group means

Ο ΠΙΝΑΚΑΣ 9 δείχνει ότι οι prior probabilities είναι ίσες, 0.5 η καθεμία, όπως άλλωστε τις είχαμε επιλέξει από την αρχή. Δεν θα αναφερθούμε ξανά στις prior probabilities καθώς εξαντλήσαμε το θέμα στην ενότητα που αναφέραμε τις υποθέσεις της discriminant analysis.

ΠΙΝΑΚΑΣ9 Prior Probabilities for Groups			
Y	Prior	Cases Used in Analysis	
		Unweighted	Weighted
0	0.5	42	42
1	0.5	42	42
Total	1	84	84

Ο ΠΙΝΑΚΑΣ 10 εμφανίζει τις πραγματικές και τις προβλεπόμενες από το μοντέλο κατατάξεις των 84 παρατηρήσεων, καθώς και τις πιθανότητες με τις οποίες αυτές κατατάσσονται. Αν προσέξουμε τις στήλες θα δούμε ότι εμφανίζονται το πραγματικό γκρουπ, το προβλεπόμενο γκρουπ, οι posterior πιθανότητες, το τετράγωνο της απόστασης Mahalanobis και τα discriminant scores.

Οι πρώτες δύο στήλες περιέχουν για τις 84 παρατηρήσεις την πραγματική τους ένταξη στο ένα από τα δύο γκρουπ και την προβλεπόμενη από το μοντέλο. Μας δίνουν τη δυνατότητα να συγκρίνουμε το πραγματικό με το προβλεπόμενο αποτέλεσμα και να αποκτήσουμε μια αίσθηση της ακρίβειας με την οποία προβλέπει το μοντέλο. Πρόκειται ουσιαστικά για το αντίστοιχο τεστ που είχε γίνει στο αρχικό δείγμα από το οποίο προήλθε το μοντέλο, στα probit και logit.

Η πιθανότητα με τον τίτλο $P(G=g | D=d)$ στο τμήμα του Highest Group είναι η posterior probability να ανήκει μια παρατήρηση στο προβλεπόμενο γκρουπ. Η ίδια πιθανότητα στον τομέα του Second Highest Group δείχνει τα αποτελέσματα για το αμέσως επόμενο πιθανότερο γκρουπ.

Η πιθανότητα με τίτλο $P(D>d | G=g)$ στο Highest Group, γνωστή και ως δεσμευμένη πιθανότητα, είναι η πιθανότητα που αντιστοιχεί στο παρατηρημένο σκορ που κατατάσσει στο πιο πιθανό γκρουπ.

Η squared Mahalanobis distance to centroids είναι ένα μέτρο της απόστασης των περιπτώσεων από τον μέσο του γκρουπ. Πρόκειται ουσιαστικά για το τετράγωνο της διαφοράς ανάμεσα στους δύο μέσους, της συνάρτησης της συγκεκριμένης παρατήρησης και του γκρουπ. Όσο πιο μεγάλο είναι, τόσο πιο μεγάλη η απόσταση και επομένως η συγκεκριμένη παρατήρηση πρέπει να αποκλειστεί από αυτό το γκρουπ.

Τα discriminant scores που εμφανίζονται στην τελευταία στήλη, υπολογίζονται πολλαπλασιάζοντας τους non-standardized discriminant coefficients με τις τιμές των 3 ανεξάρτητων μεταβλητών, προσθέτοντας αυτά τα γινόμενα και προσθέτοντας τέλος και τη σταθερά.

Ας πάρουμε όμως μία συγκεκριμένη περίπτωση για να μπορέσουμε να κατανοήσουμε αυτόν τον πίνακα. Έστω ότι μελετάμε την πρώτη γραμμή, δηλαδή την παρατήρηση 1: η δεύτερη και η τρίτη στήλη μας δείχνουν ότι η συγκεκριμένη παρατήρηση-εταιρία ανήκει στο γκρουπ 1 των υγιών και το μοντέλο την κατατάσσει επίσης στο γκρουπ 1, άρα προβλέπει σωστά.

Η τέταρτη στήλη δείχνει ότι το μοντέλο δίνει το συγκεκριμένο discriminant score της τάξεως του 2.210 (τελευταία στήλη) με πιθανότητα 0.148, όχι με ιδιαίτερα μεγάλη πιθανότητα δηλαδή.

Η πέμπτη στήλη μας δίνει τους βαθμούς ελευθερίας με τους οποίους συνδέεται αυτή η πιθανότητα.

Η έκτη στήλη μας δίνει την πιθανότητα με την οποία προβλέπει το μοντέλο ότι η συγκεκριμένη εταιρία ανήκει στο γκρουπ 1. Παρατηρούμε ότι η πιθανότητα για την παρατήρηση 1 είναι πολύ μεγάλη, κοντά στο 1, κάτι που συμβαδίζει και με την ορθότητα της πρόβλεψης.

Η έβδομη στήλη δείχνει τη squared Mahalanobis distance η οποία εμφανίζεται μικρή και άρα η παρατήρηση πρέπει να ενταχθεί στο συγκεκριμένο γκρουπ.

Η όγδοη στήλη δείχνει την πρόβλεψη για την κατάταξη στο αμέσως πιθανότερο γκρουπ. Το αμέσως πιθανότερο γκρουπ είναι φυσικά το γκρουπ 0, αφού δύο συνολικά γκρουπ έχουμε.

Η ένατη στήλη δίνει την πιθανότητα με την οποία κατατάσσεται αυτή η παρατήρηση στο γκρουπ 0, η οποία είναι μόνο 0.033, κάτι που ήταν αναμενόμενο, αφού η παρατήρηση έχει ήδη καταταχθεί με πολύ μεγάλη πιθανότητα στο άλλο γκρουπ.

Η δέκατη στήλη δείχνει και πάλι τη squared Mahalanobis distance, που εδώ εμφανίζεται πολύ μεγάλη και επομένως η παρατήρηση πρέπει να αποκλειστεί από το γκρουπ 0.

Όλα λοιπόν τα μέτρα υποδεικνύουν την κατάταξη της παρατήρησης στο γκρουπ 1.

Κατά τον ίδιο τρόπο γίνεται η ανάλυση και για όλες τις υπόλοιπες περιπτώσεις. Ενδιαφέρον παρουσιάζει η παρατήρηση των λανθασμένων προβλέψεων καθώς εκεί τα παραπάνω μέτρα δεν παρουσιάζονται τόσο ξεκάθαρα, γι' αυτό και το πρόγραμμα κάνει λάθος.

Το δεύτερο μέρος του πίνακα παρουσιάζει ακριβώς τα ίδια αποτελέσματα όχι όμως για το αρχικό δείγμα, αλλά για ένα cross-validated δείγμα, που σημαίνει ότι για κάθε περίπτωση υπολογίζονται όλα τα παραπάνω μέτρα με βάση όλο το υπόλοιπο δείγμα εκτός της συγκεκριμένης περίπτωσης. Τα αποτελέσματα για το δεύτερο αυτό δείγμα δεν διαφέρουν ιδιαίτερα από τα πρώτα, επιβεβαιώνοντας έτσι το μέγεθος της ακρίβειας πρόβλεψης του μοντέλου.

ΠΙΝΑΚΑΣ 10										
Casewise Statistics										
Original sample			Highest Group				Second Highest Group			Discrim. Scores
Case Num.	Act. Group	Pred. Group	P(D>d G=g)		P(G=g D=d)	Sq. Mahalanobis Dist.	Group	P(G=g D=d)	Sq. Mahalanobis Dist.	Function 1
			p	df						
1	1	1	0.148	1	0.967	2.089	0	0.033	8.852	2.210
2	0	0	0.464	1	0.908	0.535	1	0.092	5.115	-1.497
3	1	0	0.893	1	0.724	0.018	1	0.276	1.946	-0.630
4	0	1	0.508	1	0.899	0.438	0	0.101	4.804	1.427
5	1	1	0.767	1	0.835	0.088	0	0.165	3.337	1.062
6	0	1	0.517	1	0.545	0.419	0	0.455	0.779	0.117
7	1	1	0.666	1	0.862	0.187	0	0.138	3.850	1.197
8	0	0	0.674	1	0.629	0.177	1	0.371	1.230	-0.344
9	1	1	0.677	1	0.630	0.173	0	0.370	1.240	0.349
10	0	0	0.475	1	0.906	0.510	1	0.094	5.036	-1.479
11	1	1	0.897	1	0.797	0.017	0	0.203	2.755	0.895
12	0	1	0.609	1	0.596	0.261	0	0.404	1.038	0.254
13	1	1	0.526	1	0.550	0.402	0	0.450	0.803	0.131
14	0	1	0.582	1	0.582	0.302	0	0.418	0.960	0.215
15	1	1	0.185	1	0.961	1.757	0	0.039	8.153	2.090
16	0	0	0.680	1	0.632	0.170	1	0.368	1.249	-0.353
17	1	1	0.957	1	0.748	0.003	0	0.252	2.178	0.711
18	0	0	0.580	1	0.580	0.305	1	0.420	0.955	-0.212
19	1	1	0.656	1	0.620	0.198	0	0.380	1.176	0.320
20	0	0	0.007	1	0.995	7.199	1	0.005	17.750	-3.448
21	1	1	0.723	1	0.652	0.125	0	0.348	1.383	0.411
22	0	1	0.756	1	0.838	0.097	0	0.162	3.389	1.076
23	1	0	0.558	1	0.568	0.343	1	0.432	0.891	-0.179
24	0	0	0.849	1	0.812	0.036	1	0.188	2.961	-0.956
25	1	1	0.136	1	0.969	2.226	0	0.031	9.132	2.257
26	0	0	0.045	1	0.986	4.005	1	0.014	12.469	-2.766

27	1	1	0.865	1	0.807	0.029	0	0.193	2.891	0.935
28	0	0	0.589	1	0.880	0.291	1	0.120	4.283	-1.305
29	1	1	0.001	1	0.998	11.148	0	0.002	23.705	4.104
30	0	0	0.857	1	0.710	0.032	1	0.290	1.822	-0.585
31	1	1	0.173	1	0.963	1.859	0	0.037	8.372	2.128
32	0	0	0.272	1	0.945	1.205	1	0.055	6.903	-1.862
33	1	1	0.377	1	0.926	0.781	0	0.074	5.827	1.649
34	0	0	0.811	1	0.691	0.057	1	0.309	1.666	-0.526
35	1	1	0.997	1	0.762	0.000	0	0.238	2.331	0.762
36	0	0	0.707	1	0.851	0.142	1	0.149	3.633	-1.141
37	1	0	0.669	1	0.626	0.182	1	0.374	1.216	-0.338
38	0	0	0.628	1	0.606	0.235	1	0.394	1.092	-0.280
39	1	1	0.487	1	0.903	0.484	0	0.097	4.952	1.460
40	0	0	1.000	1	0.763	0.000	1	0.237	2.340	-0.765
41	1	1	0.804	1	0.825	0.061	0	0.175	3.160	1.013
42	0	0	0.156	1	0.966	2.014	1	0.034	8.696	-2.184

ΠΙΝΑΚΑΣ 10			Casewise Statistics							
Original sample			Highest Group				Second Highest Group			Discrim. Scores
Case Num.	Act. Group	Pred. Group	P(D>d G=g)		P(G=g D=d)	Sq. Mahalanobis Dist.	Group	P(G=g D=d)	Sq. Mahalanobis Dist.	Function 1
			p	df						
43	1	0	0.496	1	0.532	0.464	1	0.468	0.720	-0.084
44	0	0	0.654	1	0.865	0.200	1	0.135	3.911	-1.213
45	1	1	0.579	1	0.580	0.308	0	0.420	0.950	0.210
46	0	0	0.716	1	0.649	0.132	1	0.351	1.361	-0.402
47	1	0	0.943	1	0.743	0.005	1	0.257	2.128	-0.694
48	0	0	0.504	1	0.537	0.447	1	0.463	0.741	-0.096
49	1	1	0.671	1	0.627	0.180	0	0.373	1.223	0.341
50	0	0	0.438	1	0.914	0.602	1	0.086	5.318	-1.541
51	1	1	0.616	1	0.874	0.252	0	0.126	4.127	1.267

52	0	0	0.882	1	0.720	0.022	1	0.280	1.907	-0.616
53	1	0	0.807	1	0.689	0.060	1	0.311	1.652	-0.520
54	0	1	0.680	1	0.858	0.170	0	0.142	3.773	1.177
55	1	0	0.530	1	0.552	0.393	1	0.448	0.815	-0.138
56	0	0	0.897	1	0.797	0.017	1	0.203	2.755	-0.895
57	1	1	0.666	1	0.862	0.187	0	0.138	3.850	1.197
58	0	0	0.682	1	0.858	0.167	1	0.142	3.760	-1.174
59	1	1	0.421	1	0.917	0.646	0	0.083	5.447	1.569
60	0	0	0.657	1	0.864	0.198	1	0.136	3.898	-1.209
61	1	1	0.552	1	0.889	0.353	0	0.111	4.512	1.359
62	0	0	0.838	1	0.702	0.042	1	0.298	1.756	-0.560
63	1	1	0.938	1	0.784	0.006	0	0.216	2.585	0.843
64	0	0	0.831	1	0.699	0.046	1	0.301	1.732	-0.551
65	1	1	0.949	1	0.745	0.004	0	0.255	2.149	0.701
66	0	0	0.469	1	0.515	0.525	1	0.485	0.649	-0.040
67	1	1	0.485	1	0.526	0.487	0	0.474	0.693	0.067
68	0	0	0.816	1	0.693	0.054	1	0.307	1.681	-0.532
69	1	0	0.699	1	0.853	0.149	1	0.147	3.671	-1.151
70	0	0	0.749	1	0.840	0.102	1	0.160	3.420	-1.084
71	1	1	0.602	1	0.877	0.272	0	0.123	4.206	1.286
72	0	0	0.700	1	0.641	0.149	1	0.359	1.309	-0.379
73	1	1	0.656	1	0.620	0.198	0	0.380	1.176	0.320
74	0	0	0.114	1	0.973	2.497	1	0.027	9.673	-2.345
75	1	1	0.227	1	0.953	1.457	0	0.047	7.491	1.972
76	0	0	0.708	1	0.645	0.140	1	0.355	1.336	-0.391
77	1	0	0.502	1	0.535	0.452	1	0.465	0.736	-0.093
78	0	1	0.461	1	0.511	0.543	0	0.489	0.629	0.028
79	1	0	0.502	1	0.535	0.452	1	0.465	0.736	-0.093
80	0	0	0.587	1	0.584	0.295	1	0.416	0.973	-0.221
81	1	0	0.629	1	0.606	0.234	1	0.394	1.096	-0.282

82	0	0	0.421	1	0.917	0.647	1	0.083	5.450	-1.570
83	1	1	0.455	1	0.910	0.559	0	0.090	5.187	1.513
84	0	0	0.257	1	0.948	1.287	1	0.052	7.098	-1.899

ΠΙΝΑΚΑΣ 10			Casewise Statistics							
Cross validated sample			Highest Group				Second Highest Group			Discrim. Scores
Case Num.	Act. Group	Pred. Group	P(D>d G=g)		P(G=g D=d)	Sq. Mahalanobis Dist.	Group	P(G=g D=d)	Sq. Mahalanobis Dist.	Function 1
			p	df						
1	1	1	0.275	3	0.967	3.877	1	0.033	10.642	
2	0	0	0.897	3	0.906	0.596	2	0.094	5.123	
3	1	0	0.917	3	0.741	0.509	2	0.259	2.611	
4	0	1	0.090	3	0.945	6.483	1	0.055	12.182	
5	1	1	0.426	3	0.826	2.787	1	0.174	5.898	
6	0	1	0.568	3	0.566	2.021	1	0.434	2.556	
7	1	1	0.820	3	0.858	0.921	1	0.142	4.516	
8	0	0	0.654	3	0.616	1.624	2	0.384	2.568	
9	1	1	0.531	3	0.613	2.207	1	0.387	3.129	
10	0	0	0.824	3	0.903	0.907	2	0.097	5.374	
11	1	1	0.834	3	0.792	0.866	1	0.208	3.535	
12	0	1	0.912	3	0.605	0.533	1	0.395	1.386	
13	1	1	0.850	3	0.544	0.800	1	0.456	1.152	
14	0	1	0.930	3	0.589	0.449	1	0.411	1.168	
15	1	1	0.358	3	0.960	3.224	1	0.040	9.599	
16	0	0	0.210	3	0.598	4.524	2	0.402	5.315	
17	1	1	0.988	3	0.745	0.130	1	0.255	2.272	
18	0	0	0.861	3	0.574	0.750	2	0.426	1.350	
19	1	1	0.924	3	0.616	0.478	1	0.384	1.419	
20	0	0	0.000	3	0.997	21.025	2	0.003	32.773	

21	1	1	0.212	3	0.620	4.503	1	0.380	5.480	
22	0	1	0.451	3	0.875	2.639	1	0.125	6.531	
23	1	0	0.947	3	0.574	0.366	2	0.426	0.958	
24	0	0	0.941	3	0.808	0.397	2	0.192	3.271	
25	1	1	0.018	3	0.970	10.125	1	0.030	17.093	
26	0	0	0.154	3	0.987	5.259	2	0.013	13.878	
27	1	1	0.712	3	0.800	1.372	1	0.200	4.144	
28	0	0	0.918	3	0.878	0.504	2	0.122	4.442	
29	1	1	0.000	3	0.999	23.134	1	0.001	37.674	
30	0	0	0.635	3	0.698	1.710	2	0.302	3.388	
31	1	1	0.001	3	0.963	16.089	1	0.037	22.629	
32	0	0	0.661	3	0.944	1.592	2	0.056	7.245	
33	1	1	0.727	3	0.924	1.309	1	0.076	6.296	
34	0	0	0.949	3	0.687	0.359	2	0.313	1.930	
35	1	1	0.343	3	0.745	3.333	1	0.255	5.478	
36	0	0	0.968	3	0.848	0.257	2	0.152	3.702	
37	1	0	0.554	3	0.655	2.090	2	0.345	3.368	
38	0	0	0.952	3	0.603	0.341	2	0.397	1.173	
39	1	1	0.588	3	0.900	1.925	1	0.100	6.311	
40	0	0	0.661	3	0.754	1.592	2	0.246	3.829	
41	1	1	0.967	3	0.821	0.260	1	0.179	3.313	
42	0	0	0.520	3	0.966	2.261	2	0.034	8.932	

ΠΙΝΑΚΑΣ 10			Casewise Statistics							
Cross validated sample			Highest Group				Second Highest Group			Discrim. Scores
Case Num.	Act. Group	Pred. Group	P(D>d G=g)		P(G=g D=d)	Sq. Mahalanobis Dist.	Group	P(G=g D=d)	Sq. Mahalanobis Dist.	Function 1
			p	df						
43	1	0	0.910	3	0.537	0.542	2	0.463	0.839	
44	0	0	0.847	3	0.861	0.812	2	0.139	4.457	

45	1	1	0.732	3	0.569	1.289	1	0.431	1.846	
46	0	0	0.521	3	0.632	2.256	2	0.368	3.338	
47	1	0	0.962	3	0.759	0.287	2	0.241	2.579	
48	0	0	0.720	3	0.526	1.338	2	0.474	1.546	
49	1	0	0.001	3	0.509	17.655	2	0.491	17.727	
50	0	0	0.855	3	0.911	0.776	2	0.089	5.437	
51	1	1	0.622	3	0.869	1.769	1	0.131	5.554	
52	0	0	0.867	3	0.713	0.728	2	0.287	2.553	
53	1	0	0.000	3	0.861	19.140	2	0.139	22.791	
54	0	1	0.277	3	0.901	3.858	1	0.099	8.271	
55	1	0	0.918	3	0.558	0.502	2	0.442	0.972	
56	0	0	0.207	3	0.778	4.556	2	0.222	7.069	
57	1	1	0.820	3	0.858	0.921	1	0.142	4.516	
58	0	0	0.928	3	0.854	0.460	2	0.146	3.999	
59	1	1	0.465	3	0.914	2.559	1	0.086	7.274	
60	0	0	0.931	3	0.861	0.445	2	0.139	4.093	
61	1	1	0.857	3	0.886	0.768	1	0.114	4.868	
62	0	0	0.971	3	0.699	0.240	2	0.301	1.922	
63	1	1	0.195	3	0.763	4.704	1	0.237	7.044	
64	0	0	0.836	3	0.692	0.855	2	0.308	2.476	
65	1	1	0.238	3	0.722	4.231	1	0.278	6.140	
66	0	1	0.356	3	0.513	3.240	1	0.487	3.348	
67	1	1	0.662	3	0.513	1.589	1	0.487	1.691	
68	0	0	0.597	3	0.680	1.881	2	0.320	3.385	
69	1	0	0.916	3	0.874	0.512	2	0.126	4.392	
70	0	0	0.985	3	0.837	0.152	2	0.163	3.428	
71	1	1	0.836	3	0.874	0.857	1	0.126	4.727	
72	0	0	0.485	3	0.623	2.445	2	0.377	3.446	
73	1	1	0.924	3	0.616	0.478	1	0.384	1.419	
74	0	0	0.070	3	0.974	7.048	2	0.026	14.289	

75	1	1	0.574	3	0.953	1.992	1	0.047	7.989	
76	0	0	0.941	3	0.641	0.397	2	0.359	1.558	
77	1	0	0.883	3	0.542	0.660	2	0.458	0.997	
78	0	1	0.861	3	0.517	0.754	1	0.483	0.888	
79	1	0	0.883	3	0.542	0.660	2	0.458	0.997	
80	0	0	0.696	3	0.572	1.442	2	0.428	2.024	
81	1	0	0.838	3	0.619	0.849	2	0.381	1.824	
82	0	0	0.000	3	0.902	18.447	2	0.098	22.886	
83	1	1	0.554	3	0.907	2.089	1	0.093	6.636	
84	0	0	0.662	3	0.947	1.590	2	0.053	7.359	

For the original data, squared Mahalanobis distance is based on canonical functions.
For the cross-validated data, squared Mahalanobis distance is based on observations.

■ Misclassified case

a. Cross validation is done only for those cases in the analysis. In cross validation, each case is classified by the functions derived from all cases other than that case.

Ο ΠΙΝΑΚΑΣ11 τέλος, παρουσιάζει τα τελικά αποτελέσματα ακρίβειας για το μοντέλο. Το αρχικό δείγμα έχει ακρίβεια 83.33% με 35 στις 42 σωστές προβλέψεις, στην πρόβλεψη των προβληματικών εταιριών και 73.81% , με 31 στις 42 σωστές προβλέψεις, στην πρόβλεψη των υγιών. Στο cross-validated δείγμα παρουσιάζεται μια ελαφρώς μικρότερη ακρίβεια με την ακρίβεια των προβληματικών στο 80.95% και των υγιών στο 71.43%.

Παρατηρούμε ότι και εδώ υπερέχει η ακρίβεια πρόβλεψης των προβληματικών και η μείωση του σφάλματος τύπου I, κάτι που συναντήσαμε και στα προηγούμενα μοντέλα probit και logit.

Η γενική ακρίβεια του μοντέλου μας βρίσκεται στο επίπεδο του 78.6%, επίπεδο που θεωρείται αρκετά ικανοποιητικό, δεδομένων πάντα των στοιχείων που διαθέταμε. Η ακρίβεια βέβαια του μοντέλου μας πρέπει να ελεγχθεί και με κάποια τεστ ανάλογα εκείνων που διεξήγαμε στο δύο προηγούμενα μοντέλα, ώστε να μπορέσουμε να καταλήξουμε σε κάποια πιο σίγουρα συμπεράσματα για αυτήν.

ΠΙΝΑΚΑΣ11		Classification Results			
		Y	Predicted Group Membership		Total
			0	1	
Original	Count	0	35	7	42
		1	11	31	42
	%	0	83.33	16.67	100
		1	26.19	73.81	100
Cross-validated	Count	0	34	8	42
		1	12	30	42
	%	0	80.95	19.05	100
		1	28.57	71.43	100

78.6% of original grouped cases correctly classified.

76.2% of cross-validated grouped cases correctly classified.

Τεστ 1: Αρχικό δείγμα 2yrs prior

Χρησιμοποιώντας ως βάση τα δεδομένα του αρχικού πάλι δείγματος αλλά για 2 έτη πριν την καθυστέρηση, εξετάσαμε την ακρίβεια πρόβλεψης του μοντέλου. Εισάγαμε δηλαδή ως δεδομένα τους 3 δείκτες των 54 πια εταιριών, τις παραμέτρους του μοντέλου μας που είναι πια γνωστές και επιλύσαμε ως προς το άγνωστο πλέον y . Στη συνέχεια,

συγκρίναμε τα εκτιμημένα αυτά y με τα πραγματικά και διαπιστώσαμε το ποσοστό ακρίβειας του μοντέλου μας σε αυτό το δείγμα. Παρακάτω παρατίθεται ο πίνακας των αποτελεσμάτων:

test1			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 54			
Prediction Evaluation (success cutoff C = 0)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)≤C	25	9	34
P(Dep=1)>C	2	18	20
Total	27	27	54
Correct	25	18	43
% Correct	92.59	66.67	79.63
% Incorrect	7.41	33.33	20.37

Παρατηρούμε ότι το μοντέλο παρουσιάζει ακρίβεια 79.63%, ποσοστό ακόμη μεγαλύτερο από αυτό του αρχικού δείγματος. Αναμέναμε βέβαια αρκετά μεγάλη ακρίβεια καθώς ο έλεγχος έγινε στο ίδιο δείγμα από το οποίο προήλθε το μοντέλο, αλλά το γεγονός ότι οι παρατηρήσεις αυτού του δείγματος αφορούσαν 2 έτη πριν την καθυστέρηση θα έπρεπε να συνεπάγεται μικρότερη και όχι μεγαλύτερη ακρίβεια από αυτή που μας έδωσαν τα στοιχεία για το 1 έτος πριν. Παρατηρούμε επίσης ότι η ακρίβεια είναι μεγαλύτερη στην πρόβλεψη και πάλι των προβληματικών, γεγονός ευνοϊκό.

Θεωρητικά, το **cutoff score** του y είναι ο μέσος όρος των group centroids, των μέσων δηλαδή των δύο γκρουπ. Οι μέσοι αυτοί υπολογίστηκαν όπως είδαμε παραπάνω σε -0.765 για το γκρουπ των προβληματικών και 0.765 για το γκρουπ των υγιών. Ο μέσος όρος λοιπόν αυτών των δύο είναι το 0 και αυτό ακριβώς το score πήραμε, όπως αναγράφεται και στον πίνακα. Επειδή όμως παρατηρήσαμε ότι κάποιες περιπτώσεις επιχειρήσεων κατατάσσονταν οριακά λάθος, π.χ. μια προβληματική που παρουσίαζε σκορ 0.05 κατατασσόταν ως υγιής, μεταβάλλαμε κατά ένα δεκαδικό ψηφίο προς τα πάνω το σκορ και το κάναμε 0.1. Τα αποτελέσματα παρατίθενται στον επόμενο πίνακα:

test1			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 54			
Prediction Evaluation (success cutoff C = 0.1)			

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	23	5	28
P(Dep=1)>C	4	22	26
Total	27	27	54
Correct	23	22	45
% Correct	85.19	81.48	83.33
% Incorrect	14.81	18.52	16.67

Παρατηρούμε ότι η αύξηση του score έχει ως αποτέλεσμα την αύξηση της συνολικής ακρίβειας και την αύξηση της επιμέρους των υγιών με αυτή των προβληματικών όμως πάλι να υπερέχει. Παρ' όλ' αυτά, η ακρίβεια στις προβληματικές μειώνεται. Αυτά τα δύο αποτελέσματα εξηγούνται ως εξής: όταν το σκορ ήταν 0, οι υγιείς ή προβληματικές εταιρίες που παρουσίαζαν σκορ πάνω από 0 κατατάσσονταν ως υγιείς. Τώρα που το σκορ είναι 0.1, οι εταιρίες προβληματικές ή υγιείς που παρουσίασαν σκορ μεταξύ 0 και 0.1 κατατάσσονται πλέον ως προβληματικές. Αυτό σημαίνει ότι κατατάσσονται πλέον ως προβληματικές περισσότερες εταιρίες από ό,τι πρέπει και επομένως η ακρίβεια στην πρόβλεψη των προβληματικών μειώνεται. Αντίστοιχα, το νέο σκορ βγάζει ως υγιείς λιγότερες από ό,τι ήταν πριν, γεγονός που φαίνεται να συμβαδίζει με την πραγματικότητα, και άρα αυξάνει την ακρίβεια των υγιών. Γενικά, όσο το σκορ πλησιάζει το 1 τόσο περισσότερο θα αυξάνεται η ακρίβεια στην πρόβλεψη των υγιών, μέχρι ένα σημείο, όπου και οι δύο επιμέρους ακρίβειες θα αρχίσουν να μειώνονται. Γενικά, το σφάλμα τύπου I μπορεί να μειωθεί πάντα εις βάρος του σφάλματος τύπου II, δηλαδή η μία ακρίβεια αυξάνεται σε βάρος της άλλης. Το ζητούμενο είναι να βρεθεί εκείνο το σκορ, όπου και οι δύο ακρίβειες είναι όσο το δυνατόν μεγαλύτερες και επομένως η συνολική ακρίβεια γίνεται όσο το δυνατόν μεγαλύτερη. Επίσης, μας ενδιαφέρει το σφάλμα τύπου I να είναι μικρότερο από το τύπου II. Πιστεύουμε ότι το 0.1 είναι το σκορ που επιτυγχάνει όλα αυτά συγχρόνως.

Τεστ 2: Αρχικό δείγμα 1+ 2yrs prior

Με δεδομένα όλα τα στοιχεία του αρχικού δείγματος, και του 1 και των 2 ετών πριν την καθυστέρηση, κάναμε τον ίδιο έλεγχο και διαπιστώσαμε μια μικρότερη ακρίβεια αλλά ακόμη ικανοποιητική. Η ακρίβεια αυτή προσεγγίζει την ακρίβεια του original δείγματος και είναι φυσικό αφού περιέχεται και αυτό στον έλεγχο. Και πάλι η ακρίβεια κατάταξης των προβληματικών εμφανίζεται αισθητά μεγαλύτερη.

test2			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 138			
Prediction Evaluation (success cutoff C = 0)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	62	25	87
P(Dep=1)>C	7	44	51
Total	69	69	138
Correct	62	44	106
% Correct	89.86	63.77	76.81
% Incorrect	10.14	36.23	23.19

Στη συνέχεια παρατίθενται τα αποτελέσματα του ίδιου ελέγχου αλλά με διαφορετικό πάλι cutoff score = 0.1.

test2			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 138			
Prediction Evaluation (success cutoff C = 0.1)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	58	18	76
P(Dep=1)>C	11	51	62
Total	69	69	138
Correct	58	51	109
% Correct	84.06	73.91	78.99
% Incorrect	15.94	26.09	21.01

Η συνολική ακρίβεια και εδώ αυξάνεται με την αύξηση του σκορ, αλλά παρατηρούμε ότι η ακρίβεια των προβληματικών μειώνεται, ενώ των υγιών αυξάνεται, με την πρώτη όμως να παραμένει πάντα μεγαλύτερη.

Τεστ 3: 2^ο δείγμα 2yrs prior

Στη συνέχεια διεξάγουμε ένα τεστ της ίδιας μορφής σε στοιχεία όμως διαφορετικού δείγματος από το αρχικό, στο δεύτερο δείγμα που είχαμε δημιουργήσει των 42 παρατηρήσεων.

test3			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 18			
Prediction Evaluation (success cutoff C = 0)			
	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	6	2	8
P(Dep=1)>C	3	7	10
Total	9	9	18
Correct	6	7	13
% Correct	66.67	77.78	72.22
% Incorrect	33.33	22.22	27.78

Η συνολική ακρίβεια μειώνεται, κάτι που όμως ήταν αναμενόμενο, καθώς όλοι οι προηγούμενοι έλεγχοι είχαν γίνει στο αρχικό δείγμα από το οποίο είχε προέλθει η discriminant function. Τώρα που χρησιμοποιούμε διαφορετικό δείγμα, είναι λογικό η ακρίβεια να μην είναι σε τόσο υψηλά επίπεδα. Παρ' όλ' αυτά, εξακολουθεί να βρίσκεται σε ικανοποιητικά επίπεδα. Αυτό που αξίζει να επισημάνουμε είναι ότι η ακρίβεια της πρόβλεψης των προβληματικών παρουσιάζεται για πρώτη φορά μικρότερη από των υγιών. Οφείλουμε όμως να υπενθυμίσουμε ότι το συγκεκριμένο δείγμα δεν θεωρείται ιδιαίτερα κατάλληλο για την εξαγωγή συμπερασμάτων καθώς περιέχει μόνο 18 παρατηρήσεις. Στη συνέχεια αυξάνουμε το cutoff score για να διαπιστώσουμε ότι η συνολική ακρίβεια αυξάνεται σε μεγάλο βαθμό, με αυτή των υγιών να φτάνει στο 100%. Πάλι όμως πρέπει να πούμε όπως και στο probit και logit, που είχε συμβεί το ίδιο για τις προβληματικές, ότι ήταν απλά τυχαίο το ότι σε ένα τόσο μικρό δείγμα 18 παρατηρήσεων δεν προέκυψε καμία λάθος πρόβλεψη για τις υγιείς. Παρ' όλ' αυτά, ο συγκεκριμένος έλεγχος μας δίνει μια μικρή ένδειξη ότι το μοντέλο μπορεί να προβλέψει με ικανοποιητική ακρίβεια και περιπτώσεις άλλες από αυτές του αρχικού δείγματος. Μεγαλύτερη επιβεβαίωση όμως για αυτό το γεγονός θα μας δώσει το επόμενο τεστ.

test3

Dependent Variable: Y

Method: Discriminant Analysis

Sample: 1 18

Prediction Evaluation (success cutoff C = 0.1)

	Estimated Equation		
	Dep=0	Dep=1	Total
P(Dep=1)≤C	6	0	6
P(Dep=1)>C	3	9	12
Total	9	9	18
Correct	6	9	15
% Correct	66.67	100.00	83.33
% Incorrect	33.33	0.00	16.67

Τεστ 4: 2^ο δείγμα 1yr prior

Στο ίδιο αυτό δεύτερο δείγμα των 42 παρατηρήσεων για 1 έτος πριν την καθυστέρηση κάνουμε έλεγχο ακρίβειας και λαμβάνουμε μια μικρότερη ακρίβεια από την αρχική του μοντέλου, αλλά ακόμη ικανοποιητική της τάξεως του 73.68%. Η ακρίβεια των προβληματικών παρουσιάζεται και πάλι αρκετά μεγαλύτερη από των υγιών.

test4			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 38			
Prediction Evaluation (success cutoff C = 0)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	16	7	23
P(Dep=1)>C	3	12	15
Total	19	19	38
Correct	16	12	28
% Correct	84.21	63.16	73.68
% Incorrect	15.79	36.84	26.32

Η αύξηση του σκορ έχει και πάλι ως αποτέλεσμα την αύξηση της συνολικής ακρίβειας που επέρχεται με την αύξηση της ακρίβειας των υγιών.

test4			
Dependent Variable: Y			
Method: Discriminant Analysis			
Sample: 1 38			
Prediction Evaluation (success cutoff C = 0.1)			
Estimated Equation			
	Dep=0	Dep=1	Total
P(Dep=1)<=C	15	5	20
P(Dep=1)>C	4	14	18
Total	19	19	38
Correct	15	14	29
% Correct	78.95	73.68	76.32
% Incorrect	21.05	26.32	23.68

Ο παραπάνω έλεγχος μας δείχνει ότι το μοντέλο είναι ακριβές στις προβλέψεις του και για δείγματα άλλα εκτός του αρχικού. Με cutoff score 0.1 λαμβάνουμε μια αρκετά καλή ακρίβεια της τάξεως του 76.32% με πάντα μικρότερη την πιθανότητα σφάλματος τύπου I.

5.3.4. Γενικά συμπεράσματα για MDA

Η MDA αποδείχτηκε ένα επίσης ικανοποιητικό μοντέλο με μέση ακρίβεια περίπου 80.4%. Η ακρίβειά της μειώθηκε όταν την εφαρμόσαμε σε δείγμα διαφορετικό από το αρχικό, σε αναμενόμενα όμως πλαίσια.

Το cutoff score αποδείχτηκε να είναι 1 δεκαδικό ψηφίο μεγαλύτερο από το θεωρητικό και είναι τελικά το 0.1. Το συγκεκριμένο μοντέλο μπορεί να οριστεί και με τη δημιουργία κάποιας “gray area” ή αλλιώς “zone of ignorance”. Μία τέτοια ζώνη είναι ένα εύρος από y scores μέσα στην οποία παρατηρείται misclassification. Η ζώνη στη συγκεκριμένη περίπτωση θα μπορούσε να οριστεί ως το πολύ μικρό εύρος 0 – 0.1, το εύρος δηλαδή ανάμεσα στα δύο σκορ. Μέσα σε αυτό παρατηρήσαμε λάθος προβλέψεις για τις υγιείς κυρίως. Θα μπορούσαμε λοιπόν να ορίσουμε τις υγιείς επιχειρήσεις ως αυτές που εμφανίζουν σκορ πάνω από 0.1 και τις προβληματικές ως εκείνες που εμφανίζουν σκορ κάτω από 0. Όσο για τις εταιρίες που εντάσσονται ανάμεσα στα δύο σκορ, δηλαδή μέσα στην “zone of ignorance”, η κρίση μας πρέπει να είναι επιφυλακτική καθώς το μοντέλο στο συγκεκριμένο εύρος μπορεί να μην είναι επιτυχές.

Ένα βασικό πλεονέκτημα του μοντέλου είναι ότι μειώνει το σφάλμα τύπου I, την λάθος πρόβλεψη δηλαδή στις προβληματικές. Είδαμε ότι το μοντέλο αυτό “κράτησε” ως κατάλληλες και απαραίτητες μόνο τρεις μεταβλητές, των οποίων μάλιστα τα πρόσημα είναι ανάλογα της λογικής συμπεριφοράς τους. Η x_3 δηλαδή παρουσιάζει αρνητικό πρόσημο γιατί ασκεί αρνητική επίδραση στο y. Μία αύξησή της, μία αύξηση του Παθητικού εις βάρος του Ενεργητικού, θα οδηγήσει σε χειρότερη κατάσταση την εταιρία και άρα θα σημαίνει μικρότερο y. Οι άλλοι δύο δείκτες αντίθετα, παρουσιάζουν θετικό πρόσημο, όπως και θα έπρεπε άλλωστε, αφού μία αύξησή τους θα αυξήσει το y και μια μείωσή τους θα το μειώσει.

5.3.5. Γενικά συμπεράσματα για όλα τα μοντέλα

Από όλα τα παραπάνω προκύπτει ότι και τα τρία μοντέλα που εφαρμόσαμε παρέχουν μια ικανοποιητική πρόβλεψη του πιστωτικού κινδύνου. Τα τρία μοντέλα δεν διαφέρουν ιδιαίτερα μεταξύ τους, τόσο ως προς το θεωρητικό τους υπόβαθρο όσο και ως εκ τούτου ως προς τα αποτελέσματα.

Οφείλουμε για μια ακόμη φορά να επισημάνουμε τις αδυναμίες αυτής της ανάλυσης που προέρχονται από την ποιότητα και ποσότητα των στοιχείων. Οι δυσκολίες που αντιμετωπίσαμε ως προς τον αριθμό, την αντιστοίχιση και τα έτη για τα οποία είχαμε στοιχεία δεν μας εμπόδισαν να κάνουμε ό,τι ήταν καλύτερο δυνατό. Παρ’ όλ’ αυτά, παρά την όποια προσπάθειά μας, τα δεδομένα που τελικά χρησιμοποιήσαμε υστερούσαν τόσο σε ποιότητα, όσο κυρίως σε ποσότητα. Πρέπει να ομολογήσουμε ότι θα αισθανόμασταν πολύ πιο σίγουροι για τα αποτελέσματα της παραπάνω μελέτης και την ακρίβεια και καταλληλότητα των μοντέλων μας, αν είχαμε στη διάθεση μας

περισσότερα και καλύτερα στοιχεία τόσο στο αρχικό στάδιο της δημιουργίας των μοντέλων όσο και στο επόμενο στάδιο των στατιστικών ελέγχων που κάναμε.

Εντούτοις, το συμπέρασμα στο οποίο καταφέραμε να καταλήξουμε είναι ότι από τα τρία αυτά μοντέλα, ως αποτελεσματικότερο εμφανίζεται η discriminant analysis με ακρίβεια 80.4%. Ακολουθεί το probit με ακρίβεια 77.39% και τέλος το logit με 75.6%. Ένα τόσο υψηλό αποτέλεσμα βέβαια όπως αυτό της MDA, θα μπορούσε να μας βάλει σε υποψίες. Δυστυχώς δεν διαθέτουμε επιπλέον στοιχεία για τη διεξαγωγή παραπάνω ελέγχων ώστε να σιγουρευτούμε. Αντί όμως να λάβουμε υπόψη μας τον μέσο όρο της ακρίβειας για κάθε μοντέλο, ίσως να ήταν αντιπροσωπευτικότερο να λάβουμε τις ακρίβειες που παρουσιάζονται στους ελέγχους με διαφορετικά δείγματα. Ποσοστά της τάξεως του 76.32% για την MDA και 69.05% για τα logit και probit, ίσως να ανταποκρίνονται περισσότερο στην πραγματικότητα και την ανεπάρκεια των στοιχείων μας. Και με αυτά τα ποσοστά όμως, τα μοντέλα εξακολουθούν να χαρακτηρίζονται από ικανοποιητική ακρίβεια, με την MDA να εμφανίζεται καλύτερη.

Ένα επίσης σημαντικό πλεονέκτημα που παρουσιάζουν και τα τρία μοντέλα είναι η μείωση του σφάλματος τύπου I, δηλαδή η μεγάλη ακρίβεια πρόβλεψης πιστωτικού κινδύνου στον τομέα των προβληματικών. Αυτό το χαρακτηριστικό κάνει τα μοντέλα μας ιδιαίτερα ευαίσθητα σε προβληματικές περιπτώσεις, κάτι που είναι το πλέον επιθυμητό, καθώς μια λάθος πρόβλεψη για προβληματική εταιρία θα αποβεί περισσότερο ζημιογόνος, από μία αντίστοιχη για υγιή.

Τέλος, οφείλουμε πάλι να επισημάνουμε ότι στόχος αυτών των μοντέλων και οποιωνδήποτε μοντέλων τέτοιου τύπου, είναι να αποτελέσουν ένα απλά επιμέρους κριτήριο για τη λήψη επενδυτικών αποφάσεων και όχι τον μοναδικό οδηγό. Η ταυτόχρονη μελέτη πολλών επιμέρους κριτηρίων είναι αυτή που θα οδηγήσει στη σωστή επενδυτική απόφαση.

ΠΙΝΑΚΑΣ 1

ITERATIVE PROGRAM

logit																					
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22
1i epex met	2i epex met	3i epex met	4i epex met	5i epex met	ct coef	coef 1is epex met	coef 2is epex met	coef 3is epex met	coef 4is epex met	coef 5is epex met	p-val ct	p-val 1is epex met	p-val 2is epex met	p-val 3is epex met	p-val 4is epex met	p-val 5is epex met	correct%	p-val<0.05	P(LR stat)	McFadden R ²	H-L Statis
2	5	6	11	13	-2.009	-5.273	2.689	7.687	2.398	1.805	0.033	0.002	0.043	0.000	0.001	0.011	77.38	yes	1.12E-07	0.348	14.9
2	6	8	11	12	-0.534	-3.954	6.556	1.116	2.442	6.802	0.373	0.007	0.000	0.043	0.001	0.041	77.38	yes	9.78E-08	0.35	7.2
2	6	9	11	13	-0.354	-4.345	7.392	0.427	2.505	2.356	0.566	0.004	0.000	0.027	0.001	0.004	77.38	yes	4.26E-08	0.36	5.9
2	6	10	11	13	-0.341	-4.382	7.463	0.432	2.521	2.383	0.581	0.004	0.000	0.027	0.001	0.004	77.38	yes	4.23E-08	0.366	6
2	6	11	12	13	-0.526	-3.903	6.537	2.409	7.341	1.342	0.383	0.008	0.000	0.001	0.025	0.039	78.57	yes	8.84E-08	0.35	7.2
probit																					
2	5	6	8	11	-1.201	-3.185	1.622	4.673	0.897	1.402	0.018	0.001	0.031	0.000	0.004	0.000	76.19	yes	8.16E-08	0.35	7.4
2	5	6	11	13	-1.220	-3.084	1.654	4.583	1.361	1.025	0.017	0.001	0.029	0.000	0.000	0.005	77.38	yes	1.06E-07	0.34	12.7
2	6	8	9	11	-0.202	-2.564	4.369	1.065	0.226	1.449	0.564	0.002	0.000	0.003	0.050	0.000	78.57	yes	5.06E-08	0.36	6.58
2	6	8	11	12	-0.289	-2.343	3.943	0.665	1.399	4.078	0.402	0.004	0.000	0.041	0.000	0.044	78.57	yes	8.51E-08	0.35	6.4
2	6	9	11	13	-0.174	-2.576	4.425	0.262	1.429	1.418	0.622	0.003	0.000	0.025	0.000	0.003	77.38	yes	3.60E-08	0.36	3.47
2	6	10	11	13	-0.165	-2.598	4.468	0.266	1.436	1.435	0.642	0.003	0.000	0.025	0.000	0.003	77.38	yes	3.58E-08	0.37	3.48
2	6	11	12	13	-0.278	-2.319	3.921	1.378	4.402	0.805	0.420	0.005	0.000	0.000	0.027	0.041	77.38	yes	7.68E-08	0.35	6.17

Οι παραπάνω συνδυασμοί είναι οι συνδυασμοί στους οποίους έχουμε καταλήξει με κριτήρια τις p-values και τη συνολική ακρίβεια πρόβλεψης του κάθε συνδυασμού. Από εκεί και πέρα, μελετάμε τους καλύτερους συνδυασμούς που μας έχει δώσει το iterative program με βάση αυτά τα δύο κριτήρια και από αυτούς πλέον και με τη βοήθεια και άλλων στατιστικών κριτηρίων επιλέγουμε τον καλύτερο.

Στήλες 1- 5 : μας δίνουν τον εκάστοτε συνδυασμό των 5 μεταβλητών

Στήλες 6- 11: μας δίνουν τις παραμέτρους της σταθεράς και του εκάστοτε συνδυασμού των 5 μεταβλητών

Στήλες 12-17: μας δίνουν τις p-values της σταθεράς και των 5 διαφορετικών κάθε φορά μεταβλητών. Όταν $p\text{-value} < 0.05$ θεωρούμε ότι οι μεταβλητές μας είναι στατιστικά σημαντικές.

Στήλη 18 : μας δίνει το ποσοστό ακρίβειας του συγκεκριμένου συνδυασμού μεταβλητών

Στήλη 19 : μας ενημερώνει για το αν οι p-values όλων των μεταβλητών εκτός της σταθεράς είναι μικρότερες του 0.05.

Αν p-values όλων των μεταβλητών < 0.05 τότε έχουμε την απάντηση yes. Σε κάθε άλλη περίπτωση λαμβάνουμε την απάντηση no.

Στήλη 20 : μας δίνει την πιθανότητα του LR statistic. Όσο μικρότερη τόσο καλύτερος ο συνδυασμός των μεταβλητών, δηλαδή το μοντέλο.

Στήλη 21 : μας δίνει το McFadden R^2 . Όσο πλησιάζει προς τη μονάδα τόσο καλύτερο το μοντέλο μας.

Παρατηρούμε ότι οι p-values όλων των μεταβλητών είναι μικρότερες του 0.05, γι' αυτό και λαμβάνουμε την απάντηση yes, το p LR statistic είναι για όλους τους συνδυασμούς πάρα πολύ μικρό, το McFadden R^2 κυμαίνεται συνέχεια στο 0.35 και η ακρίβεια των διαφόρων συνδυασμών και για τις δύο μεθόδους εμφανίζεται συνέχεια μεταξύ του 76 και 78%. Οι διαφορές ανάμεσα στους διαφορετικούς συνδυασμούς είναι πραγματικά απειροελάχιστες όσο διαφορετικά και περισσότερα στατιστικά κριτήρια επιλογής και αν πάρουμε. Με βάση λοιπόν όλα τα παραπάνω και τη δική μας κρίση καταλήξαμε στους συνδυασμούς που έχουμε σημειώσει με το μωβ χρώμα. Οφείλουμε όμως να επισημάνουμε ότι και οποιοσδήποτε άλλος από τους παραπάνω συνδυασμούς θα κατέληγε σε ένα εξίσου ακριβές και αποτελεσματικό μοντέλο. Αυτό που αξίζει ίσως να παρατηρήσουμε από όλη αυτή τη μεθοδολογία είναι ότι ως χρήσιμοι για την προβλεψιμότητα των μοντέλων πιστωτικού κινδύνου εμφανίζονται οι δείκτες κερδοφορίας, ρευστότητας και αποδοτικότητας με επικρατέστερους τους x_2 , x_6 και αμέσως μετά x_{11} και x_{13} , δηλαδή αντίστοιχα : Κυκλοφορούν Ενεργητικό / Ενεργητικό, Κεφάλαιο Κίνησης / Ενεργητικό, Πωλήσεις / Ενεργητικό και Δείκτης Αποδοτικότητας Ιδίων Κεφαλαίων.

ΠΙΝΑΚ					
Variables Not in the Analysis					
ΑΣ2					
Step		Tolerance	Min. Tolerance	Sig. of F to Enter	Wilks' Lambda
0	X1	1	1	0.3494	0.9893
	X2	1	1	0.5948	0.9965
	X3	1	1	0.0160	0.9313
	X4	1	1	0.7802	0.9990

	X5	1	1	0.3615	0.9898
	X6	1	1	0.0091	0.9200
	X7	1	1	0.0002	0.8472
	X8	1	1	0.0084	0.9182
	X9	1	1	0.1448	0.9742
	X10	1	1	0.1770	0.9779
	X11	1	1	0.0024	0.8928
	X12	1	1	0.0005	0.8604
	X13	1	1	0.0100	0.9219
1	X1	1.0000	1.0000	0.3829	0.8392
	X2	0.9612	0.9612	0.8352	0.8468
	X3	0.9998	0.9998	0.0305	0.7994
	X4	0.9994	0.9994	0.8644	0.8469
	X5	0.9478	0.9478	0.9705	0.8472
	X6	0.9983	0.9983	0.0113	0.7824
	X8	0.8310	0.8310	0.2635	0.8342
	X9	0.6530	0.6530	0.3719	0.8389
	X10	0.6733	0.6733	0.3539	0.8382
	X11	0.8907	0.8907	0.0741	0.8143
	X12	0.0514	0.0514	0.6921	0.8456
	X13	0.8982	0.8982	0.1774	0.8283
2	X1	0.9793	0.9776	0.2313	0.7684
	X2	0.8985	0.8985	0.3924	0.7752
	X3	0.7554	0.7543	0.3148	0.7725
	X4	0.9727	0.9716	0.5661	0.7791
	X5	0.9265	0.9265	0.6821	0.7807
	X8	0.7435	0.7435	0.0505	0.7456
	X9	0.6529	0.6525	0.4100	0.7757
	X10	0.6730	0.6730	0.4055	0.7756
	X11	0.8159	0.8159	0.0125	0.7233
	X12	0.0503	0.0503	0.9862	0.7824
	X13	0.8034	0.8034	0.0290	0.7368
3	X1	0.9682	0.8066	0.3722	0.7160
	X2	0.6563	0.5960	0.0168	0.6725
	X3	0.5707	0.5707	0.0145	0.6703
	X4	0.9494	0.7964	0.8613	0.7231
	X5	0.8438	0.7430	0.7251	0.7222
	X8	0.7394	0.7394	0.0892	0.6972
	X9	0.6443	0.5794	0.6082	0.7209
	X10	0.6637	0.5959	0.6075	0.7209
	X12	0.0480	0.0465	0.6072	0.7209
	X13	0.7980	0.7980	0.0571	0.6907

ΠΙΝΑΚΑΣ 2		Variables Not in the Analysis			
Step		Tolerance	Min. Tolerance	Sig. of F to Enter	Wilks' Lambda
4	X1	0.9598	0.5658	0.2809	0.6603
	X2	0.2978	0.2589	0.4004	0.6642
	X4	0.8722	0.5243	0.3779	0.6636
	X5	0.8434	0.5704	0.7751	0.6696
	X8	0.5955	0.4596	0.0035	0.6005
	X9	0.5887	0.5214	0.2148	0.6571
	X10	0.6097	0.5243	0.2237	0.6576
	X12	0.0473	0.0463	0.8429	0.6699
	X13	0.6681	0.4778	0.0029	0.5979
5	X1	0.9632	0.6152	0.3146	0.6835
	X2	0.6554	0.5652	0.5441	0.6891
	X4	0.8724	0.6166	0.3542	0.6849
	X5	0.8452	0.5769	0.7136	0.6912
	X6	0.7539	0.5707	0.1105	0.6703
	X8	0.6021	0.5148	0.0049	0.6260
	X9	0.6060	0.5813	0.1291	0.6724
	X10	0.6288	0.5976	0.1333	0.6728
	X12	0.0501	0.0486	0.8550	0.6921
	X13	0.6775	0.5373	0.0044	0.6245
6	X1	0.9637	0.7168	0.3201	0.7060
	X2	0.6752	0.6125	0.3761	0.7078
	X4	0.8732	0.6620	0.3194	0.7060
	X5	0.8589	0.6558	0.8654	0.7146
	X6	0.7553	0.6014	0.1179	0.6932
	X7	0.8438	0.6166	0.1110	0.6924
	X8	0.7912	0.6252	0.0012	0.6260
	X9	0.8795	0.6510	0.6881	0.7134
	X10	0.8879	0.6562	0.6649	0.7132
	X12	0.8692	0.6360	0.1326	0.6948
	X13	0.7952	0.6218	0.0011	0.6252
7	X1	0.9549	0.6052	0.2203	0.6134
	X2	0.6227	0.4839	0.0798	0.6013
	X4	0.7884	0.6010	0.9585	0.6252
	X5	0.8307	0.6173	0.6763	0.6238
	X6	0.7435	0.5425	0.0641	0.5985
	X7	0.7189	0.5373	0.7567	0.6245
	X8	0.0284	0.0284	0.8310	0.6249
	X9	0.8370	0.5870	0.7454	0.6244
	X10	0.8420	0.5909	0.7500	0.6244
	X12	0.7541	0.5421	0.7753	0.6246