

HEC MONTRÉAL

Prévision de faillite : ratios bruts versus ratios
relatifs

Par

Hesham Abdul Menhem

Science de la gestion
Ingénierie Financière

*Mémoire présenté en vue de l'obtention
du grade de maîtrise ès science
(M. Sc.)*

Août 2014

© Hesham Abdul Menhem, 2014

Liste des Tableaux

Tableau 1 : Ratios financiers.....	11
Tableau 2 : Beaver (1966).....	13
Tableau 3 : Altman (1968).....	14
Tableau 4 : Ohlson (1980)	14
Tableau 5 : Zmijewski (1984).....	15
Tableau 6 : Chava et Jarrow (2004).....	15
Tableau 7 : Exemple de DLRSN et AT_FN	17
Tableau 8 : Secteurs économiques selon le <i>Standard Industrial Classification</i>	18
Tableau 9 : SIC versus SICH	19
Tableau 10 : Modes d'échantillonnage	21
Tableau 11 : Y x Y.....	24
Tableau 12 : Zmijewski (1984).....	29
Tableau 13 : Statistiques descriptives	29
Tableau 14 : Quartiles des distributions.....	30
Tableau 15 : Tests de signification globale.....	30
Tableau 16 : Estimation du modèle.....	30
Tableau 17 : CMMA (en %) par seuil de coupure ($c = 2$)	32
Tableau 18 : Faillites et valeur d'actif (2010-2013).....	35
Tableau 19 : Liste des 10 plus grandes faillites de 2013.....	36
Tableau 20 : Sensibilité et spécificité	37

Liste des Figures

Figure 1 : État de la situation financière	9
Figure 2 : État des résultats.....	10
Figure 3 : Exemple de décomposition de la variable SICH.....	19
Figure 4 : Point sur la courbe ROC au seuil $p^* = 0.50$	25
Figure 5 : Courbe ROC – X et Y indépendants	26
Figure 6 : Sections de notre programme informatique.....	28
Figure 7 : Courbes ROC	32
Figure 8 : Nombre de faillites par secteur d'activité (2013).....	35

Table des Matières

Liste des Tableaux	iii
Liste des Figures	iv
Remerciements.....	1
Section 1: Introduction.....	2
Section 2: Revue de littérature	4
Section 3: États financiers et ratios	9
3.1 États financiers.....	9
3.2 Ratios de Ross et al. (2010)	10
3.3 Ratios additionnels.....	13
Section 4 : Données	16
4.1 Identifiant unique	16
4.2 Statut de la compagnie	16
4.3 Date de faillite.....	17
4.4 Secteur d'activité	18
Section 5 : Choix méthodologiques	20
5.1 Échantillonnage.....	20
5.2 Benchmark et ratios relatifs	21
5.3 Modèle logistique.....	22
5.4 Indicateurs de performance.....	23
Section 6 : Investigation empirique	27
6.1 Paramètres de flexibilité	28
6.2 Résultats pour un modèle arbitraire	29
6.3 Résultats globaux	33
Section 7 : Conclusion	34
Annexe 1 : Données sur la faillite corporative.....	35
Annexe 2 : Sensibilité et spécificité.....	37
Bibliographie.....	39

Remerciements

Je remercie mon directeur de mémoire, M. Hatem Ben Ameer, pour son aide, sa patience et le savoir qu'il m'a transmis. Il a su me motiver et me pousser à améliorer continuellement toutes les facettes de ma recherche.

Je remercie également M. Mohamed Jabir, analyste de données au Laboratoire de calcul et d'exploitation des données, pour son aide précieuse lors de la consultation de bases de données.

Finalement, je souhaite remercier ma famille et mes proches pour leur support tout au long de la maîtrise. Leur aide et leur soutien ont été essentiels durant mes études.

Section 1: Introduction

La récente crise économique rappelle l'importance de la gestion de risque en finance. Les multiples faillites et mises à pied motivent l'industrie ainsi que le milieu académique à améliorer les modèles de prévision de faillite. Ce mémoire porte sur l'utilisation des ratios relatifs en prévision de faillite corporative. L'objectif de ce travail est de déterminer si les ratios relatifs permettent une meilleure prévision que les ratios bruts. À cette fin, nous comparons la performance d'une multitude de modèles en utilisant les ratios relatifs et les ratios bruts.

Ross et al. (2010) suggèrent l'usage de ratios relatifs en analyse financière plutôt que l'usage de ratios bruts. Un ratio relatif par rapport à un benchmark bien choisi renseignerait mieux sur la difficulté/santé financière d'une compagnie qu'un ratio brut. Par exemple, au lieu de considérer le ratio de fonds de roulement, nous considérons le même ratio relatif à la moyenne du secteur d'activité de la compagnie observé à la même année. Nous utilisons le modèle logistique en prévision de faillite corporative pour vérifier si les ratios relatifs renseignent mieux sur la faillite que les ratios brut.

Nous considérons les ratios, les modes d'échantillonnage, le traitement des valeurs extrêmes et les indicateurs de performance proposés par Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984), Shumway (2001), Chava et Jarrow (2004) ainsi que Ross et al. (2010). Nous proposons également plusieurs choix pour le benchmark. Nous retenons la plus grande flexibilité en programmation informatique sous SAS en permettant le croisement entre tous ces choix méthodologiques, ce qui aboutit à des centaines de modèles différents. Nous étudions la totalité des modèles, mais nous reportons les résultats d'un modèle arbitraire.

Notre analyse aboutit à des résultats mitigés. Les ratios relatifs assurent parfois une meilleure prévision, mais sont parfois dominés par les ratios bruts. Nous arrivons à la conclusion que la difficulté réside dans le choix du benchmark. Ce dernier est construit sur la base du secteur économique, qui reflète l'activité principale de la compagnie. Comme nous opérons à même la base Compustat, avec les plus grandes compagnies

boursières en Amérique du Nord, un benchmark basé sur l'activité principale n'est pas nécessairement représentatif des activités multiples de la compagnie.

Section 2: Revue de littérature

La faillite corporative a des répercussions économiques et sociales sévères. Depuis plusieurs années, les économistes et académiciens se penchent sur ce phénomène afin de mesurer le risque de faillite et ultimement prévenir son occurrence. Les conséquences de la faillite sont toutefois encore présentes aujourd'hui. BankruptcyData, une base de données gérée par la firme New Generation Research, offre des données sur les plus grandes compagnies boursières nord-américaines. Son rapport reporte 71 compagnies en faillite en 2013 avec un actif combiné de l'ordre de 43 milliards de dollars. L'Annexe 1 donne des statistiques générales concernant la faillite de grandes compagnies boursières sur la période 2010-2013.

Le risque de faillite est directement lié au problème d'allocation de crédit dans le secteur bancaire. Les modèles de prévision de faillite tentent de détecter les signes de faillite suffisamment à l'avance pour éviter les pertes en capital et en postes de travail. L'analyse des ratios comptables et financiers en prévision de faillite corporative représente un outil important en gestion de risque financier. La revue qui suit présente un survol des travaux des 50 dernières années en prévision de faillite corporative.

Les ratios comptables et financiers peuvent être classés en cinq familles, à savoir la liquidité, la structure de capital, la rentabilité, l'efficacité et le rapport au marché. La littérature suggère qu'une prévision de faillite effective sur la base des ratios comptables et financiers est tout à fait possible. Ces ratios sont alors utilisés pour discriminer entre les compagnies en faillite et les compagnies saines. La littérature dans ce domaine est riche, abondante et variée depuis les années 1960.

La prévision de faillite est discutée dans la littérature à travers trois types de modèles. Les modèles dominant sont d'ordre statistique, par exemple le modèle logistique (Ohlson 1980). Aussi, les modèles rationnels d'options jouent un rôle important, un exemple est le modèle structurel (Merton 1974). Finalement, les modèles d'apprentissage automatique prennent de plus en plus d'importance dans la littérature. Un exemple est le modèle de réseau de neurones (Bishop 1994). Dans ce qui suit, nous

concentrons la revue sur les modèles statistique puisque notre investigation empirique s'appuie sur un modèle statistique de base.

Beaver (1966) utilise les données de *Moody's Industrial Manual* sur la période 1954-1964. L'auteur considère une liste de 30 ratios comptables appartenant aux différentes familles. Il identifie les six ratios qui séparent le mieux les compagnies en faillite et les compagnies saines. Par analyse univariée, il identifie le seuil qui minimise l'erreur de classification pour chaque ratio. À la même époque, Altman (1968) utilise la même source sur la période 1946-1965. L'auteur considère quatre ratios comptables et un ratio de marché (valeur marchande sur valeur comptable). Puis, il procède à une analyse discriminante multiple qui permet d'identifier la combinaison linéaire de ratios qui discrimine au mieux entre les compagnies en faillite et les compagnies saines.

Altman poursuit son analyse avec de multiples collaborateurs (Altman et Lorriss 1976, Altman et McGough 1974, Altman, Haldeman et Narayanan 1977, Altman, Marco et Varetto 1994 et Altman et Narayanan 1997). Ces travaux sont des répliques du modèle d'Altman (1968) avec des données plus récentes. Plusieurs autres auteurs poursuivent la recherche dans la même lignée que Beaver et Altman. Deakin (1972) reprend les ratios présentés par Beaver (1966) et utilise une analyse discriminante similaire à celle d'Altman (1968) pour prédire la faillite. Wilcox (1973) propose de créer des états de difficulté/santé financière en se basant sur la valeur de ratios comptables. Dans son modèle, une compagnie passe d'un état à l'autre selon le mouvement des ratios. Blum (1974) propose de raffiner le choix des variables explicatives pour inclure des ratios dépendant du temps, comme le rendement sur une période d'investissement. Il utilise également des mesures de dispersion, comme l'écart-type des ratios, en tant que régresseurs pour ensuite passer à une analyse discriminante. Grice et Ingram (2001) reprennent le modèle d'Altman (1968) pour mesurer sa performance avec des données actuelles, sa capacité à inclure plusieurs secteurs d'activité et sa capacité à détecter des indicateurs de stress financier. Plusieurs papiers explorent les applications de l'analyse de ratios ainsi que les problèmes potentiels liés aux méthodologies des études des années 1960 (Edmister 1972, Libby 1975 et Moyer 1977).

Malgré leur popularité, ces travaux font l'objet de critiques importantes. D'abord, il existe plusieurs ratios dans chaque famille. Par exemple, la liquidité peut être mesurée par le ratio de fonds de roulement, le ratio de trésorerie ou encore le ratio de liquidité immédiate. Le choix d'une mesure par rapport à une autre ne s'appuie pas sur des critères précis. De plus, les travaux se basent sur des valeurs comptables qui ne reflètent pas nécessairement les valeurs économiques/marchande. À titre d'exemple, une dette à long terme est reportée au passif du bilan au taux exigé à la date d'émission, alors que la valeur marchande nécessite la connaissance du taux exigé à la date courante d'évaluation. Ensuite, les observations d'une même compagnie dans le temps sont corrélées, mais les modèles supposent généralement l'indépendance. Dans cette veine, les modes d'échantillonnage basés sur l'appariement, fréquemment utilisés en prévision de faillite, sont employés pour préserver l'indépendance. Pourtant, l'appariement n'est pas supporté par la théorie statistique.

Ohlson (1980) utilise les données de Compustat et les rapports obligatoires des compagnies listées en bourses (10-K) sur la période 1970-1976. L'auteur observe la faillite un certain temps avant son occurrence (deux ans avant la faillite, par exemple). Par la suite, pour chaque compagnie saine, il choisit au hasard une année d'observation. Ainsi, avec ce mode d'échantillonnage, l'hypothèse d'indépendance est acceptable. En plus des ratios comptables classiques, Ohlson ajoute des indicatrices pour mesurer la rentabilité de la compagnie sur une période de deux ans. L'auteur utilise finalement un modèle logistique pour prévoir la faillite.

Zmijewski (1984) utilise plusieurs sources de données (Compustat, CRSP, 10-K, *Capital Changes Reporter* et *Wall Street Journal Index*) sur la période 1972-1978. De plus, il interpole certaines valeurs manquantes à l'aide de valeurs renseignées. Il étudie le biais inhérent aux différentes méthodes d'échantillonnage utilisées en prévision de faillite. Il analyse l'usage de l'appariement déterministe entre compagnies en faillite et compagnies saines et l'usage de données incomplètes (données manquantes). L'auteur donne une illustration de son analyse à l'aide du modèle probit.

Shumway (2001) utilise Compustat et CRSP sur la période 1962-1992. L'auteur consulte également les sources suivantes pour compléter les données manquantes : le

Directory of Obsolete Securities et Nexis. Il interpole également certaines valeurs manquantes à l'aide de valeurs renseignées. Il utilise un modèle de durée de vie, où la durée depuis l'introduction en bourse joue un rôle clé. Shumway considère toutes les observations communes à Compustat et CRSP. Dans ce contexte, l'hypothèse d'indépendance, forcée par l'auteur, est difficile à justifier en face d'observations sérielles (même compagnie dans le temps).

Plusieurs autres auteurs introduisent le secteur d'activité dans l'analyse (Lang et Stultz 1992, Shleifer et Vishny 1992, Opler et Titman 1994, Maksimovic et Phillips 1998 et Chava et Jarrow 2004). Ces auteurs étudient les facteurs industriels, comme le cycle de vie du secteur, la demande des consommateurs et les transactions entre compagnies d'un même secteur. Ils concluent que le secteur a une influence significative sur la faillite. C'est pourquoi, durant les années 2000, l'analyse des ratios en prévision de faillite incorpore l'effet du secteur d'activité.

Chava et Jarrow (2004) reprennent les données de Compustat, CRSP et les rapports émis par les compagnies sur la période 1962-1999. Ils interpolent aussi certaines valeurs manquantes à l'aide de valeurs renseignées. En plus d'utiliser les ratios comptables, ils divisent l'ensemble des secteurs d'activité en quatre groupes et assigne une indicatrice à chacun. Les auteurs utilisent un modèle logistique pour prévoir la faillite.

Par ailleurs, la littérature discute de la performance des modèles de prévision de faillite (Sinha et May 2004, Stein 2005, Stein 2007, Chava et Jarrow 2004 et Zhou 2013). Deux indicateurs sont proposés, il s'agit du coût moyen d'une mauvaise affectation et du ROC (*Receiver Operating Characteristic*). Nous discutons ces deux indicateurs en détail plus loin dans le texte.

Depuis ces articles fondateurs, la littérature n'a cessé d'évoluer et de proposer des modèles améliorés (modèle logistique en panel, à coefficients variables, mixte, multinomial, de choix). Tous utilisent les ratios comptables et financiers pour prévoir la faillite. Barros et al. (2007) proposent un modèle logistique mixte pour prévoir la performance dans l'industrie bancaire en Europe. Hensher et al. (2007) utilisent un modèle logistique multinomial en considérant quatre états de santé (détresse) financière.

Leur analyse porte sur le risque de faillite corporative en Australie. Nam et al. (2008) utilisent un modèle de durée discret et incorpore des variables macroéconomiques dans leur prévision de faillite. Du Jardin (2010) utilise différents modèles (logistique, neurones et analyse discriminante multiple) en y appliquant plusieurs sélections de ratios comptables pour déterminer l'importance des variables explicatives. Plusieurs autres auteurs étudient ces types de modèles dans la littérature moderne (Hensher et Jones 2007, Cheng et al. 2010, Wu et al. 2010, Tsai 2013).

Nous considérons le modèle logistique classique, estimé sous l'hypothèse d'indépendance, pour vérifier si les ratios relatifs performant mieux que les ratios bruts. En effet, notre objectif est de déterminer quelles variables explicatives (ratios bruts ou ratios relatifs) offrent une meilleure prévision et non de juger entre l'efficacité des différents modèles. C'est pourquoi nous travaillons uniquement avec le modèle logistique, mais avec des variables explicatives différentes.

Notre analyse aboutit à des résultats mitigés. Les ratios relatifs assurent parfois une meilleure prévision, mais sont parfois dominés par les ratios bruts. Nous arrivons à la conclusion que la difficulté réside dans le choix du benchmark. Ce dernier est construit sur la base du secteur économique, qui reflète l'activité principale de la compagnie. Comme nous opérons à même la base Compustat, avec les plus grandes compagnies boursières en Amérique du Nord, un benchmark basé sur l'activité principale n'est pas nécessairement représentatif des activités multiples de la compagnie.

Le reste de ce mémoire est structuré comme suit. La Section 3 porte sur les états financiers et les ratios. La Section 4 traite des données et de la base Compustat. La Section 5 présente les modes d'échantillonnage, la construction du benchmark, le modèle logistique et les indicateurs de performance. La Section 6 discute de l'investigation numérique en présentant les paramètres de flexibilité et les résultats d'un modèle arbitraire. Finalement la Section 7 conclut notre travail.

Section 3: États financiers et ratios

3.1 États financiers

Les ratios comptables et financiers pour une compagnie sont construits à partir de ses états financiers. Nous considérons les ratios issus de l'état de la situation financière (le bilan), de l'état des résultats, de l'état des flux de trésorerie et du marché. L'état de la situation financière présente l'information sur la structure de capital de la compagnie, c'est-à-dire, son niveau d'actif, d'endettement et de capitaux propres. L'état des résultats contient le niveau des revenus et des dépenses encourues ainsi que le profit réalisé sur la période d'exercice. L'état des flux de trésorerie présente les fluctuations monétaires liées aux opérations, au financement et à l'investissement de la compagnie. Les deux figures suivantes illustrent les formats typiques de l'état de la situation financière et l'état des résultats.

Figure 1 : État de la situation financière

Actif à court terme	Passif à court terme
Encaisse	Dette à court terme
Stock	Comptes fournisseurs
Comptes clients	Passif à long terme
Actif immobilisé	Dette à long terme
Usines et équipement	Fonds Propres
Terrains et propriétés	Actions ordinaires
	Bénéfices non répartis
Total de l'actif	Total du passif et des fonds propres

Figure 2 : État des résultats

Revenus
- Coût des marchandises vendues
- Amortissement
- Charges et dépenses administratives
= Bénéfice avant intérêts et impôts
- Intérêts
= Bénéfice imposable
- Impôts
= Bénéfice net
Addition aux bénéfices non répartis
Dividendes

3.2 Ratios de Ross et al. (2010)

Ross et al. (2010) divise les ratios en cinq familles pour mesurer la difficulté/santé financière de la compagnie (voir Tableau 1). Nous augmentons cette liste avec les ratios de Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984) et Chava et Jarrow (2004).

Dans le Tableau 1, CT et LT signifient respectivement court terme et long terme. Le fonds de roulement représente l'actif à court terme moins le passif à court terme. Le passif représente la somme du passif à court terme et du passif à long terme. Les ventes représentent les revenus à l'état des résultats. Le bénéfice brut est égal aux revenus moins le coût des marchandises vendues et le bénéfice d'exploitation est équivalent au bénéfice avant intérêts et impôts. Le bénéfice par action est obtenu en divisant le bénéfice net par le nombre d'actions.

Tableau 1 : Ratios financiers		
Ratios	Interprétation	Définition
Liquidité		
Fonds de roulement	Capacité d'honorer les engagements à court terme avec les actifs courts termes	Actif CT / Passif CT
Trésorerie	Capacité d'honorer les engagements à court terme avec l'encaisse et l'équivalent	(Actif CT – Stock) / Passif CT
Liquidité immédiate	Capacité d'honorer les engagements à court terme avec l'encaisse uniquement	Encaisse / Passif CT
Fonds de roulement – actif	Proportion du fonds de roulements par rapport aux actifs	Fonds de roulement / Actif
Structure de capital		
Endettement total	Montant de dette par dollar d'actif	Passif / Actif
Dette - Fonds propres	Structure de financement (effet de levier)	Passif / Fonds propres
Actif - Fonds propres	Montant d'actif par dollar de fonds propres	Actif / Fonds propres
Endettement à long terme	Proportion de la dette long terme par rapport au capital disponible	Dette LT / (Dette LT + Fonds propres)
Couverture des intérêts	Capacité de payer l'intérêt sur la dette	Bénéfice avant intérêts et impôt / Intérêts
Couverture des intérêts avant amortissement	Capacité de payer l'intérêt sur la dette avec les revenus d'opération	(Bénéfice avant intérêts et impôt + Amortissement) / Intérêts
Efficacité		
Rotation des stocks	Nombre de fois que les stocks sont vendus	Coûts des marchandises vendues / Stock
Délai moyen de rotation des stocks	Nombre de jours moyen pour vendre les stocks	365 / Rotation des stocks
Rotation des comptes clients	Capacité à collecter les créances	Ventes / Comptes clients
Délai moyen de recouvrement des créances	Nombre de jours moyen pour recouvrir les comptes clients	365 / Rotation des comptes clients
Rotation du fonds de roulement	Montant de ventes par dollar de fonds de roulement	Ventes / Fonds de roulement
Rotation de l'actif	Montant de vente par dollar d'actif	Ventes / Actif
Rotation de l'actif immobilisé	Montant de vente par dollar d'actif immobilisé	Ventes / Actif immobilisé
Rentabilité		
Marge bénéficiaire nette	Montant de bénéfice net généré par dollar de ventes	Bénéfice net / Ventes
Marge bénéficiaire brute	Montant de bénéfice brut généré par dollar de ventes	Bénéfice brut / Ventes
Marge d'exploitation	Montant de bénéfice d'exploitation par dollar de ventes	Bénéfice d'exploitation / Ventes
Taux de rendement de l'actif	Montant de bénéfice net par dollar d'actif	Bénéfice net / Actif
Taux de rendement des fonds propres	Montant de bénéfice net par dollar de fonds propres	Bénéfice net / Fonds propres
Marché		
Bénéfice par action	Montant de bénéfice net par action	Bénéfice net / Nombre d'action
Cours-Bénéfice	Combien l'investisseur est prêt à payer par dollar de bénéfice (rendement futur)	Cours / Bénéfice par action
Valeur marchande - Valeur comptable	Proportion de la valeur marchande par rapport à la valeur aux livres par action	Cours / Prix aux livres de l'action

Les ratios de liquidité mesurent la capacité d'une compagnie à honorer ses engagements à court terme. Par exemple, le ratio de fonds de roulement (actif court terme sur passif court terme) indique le nombre de fois qu'une compagnie peut payer ses dettes à court terme en utilisant uniquement son actif à court terme. Plus ce ratio est élevé, plus la liquidité de la compagnie est grande.

Les ratios de structure de capital décrivent l'effet de levier (l'endettement) utilisé par la compagnie. À titre d'exemple, le ratio de dette sur fonds propres exprime le financement de l'actif. Un ratio unitaire signifie que l'actif est financé à parts égales par la dette et les fonds propres. Un ratio supérieur à 1 signifie que la compagnie utilise plus de dette pour l'acquisition d'actifs, l'effet de levier est plus élevé (l'inverse est observé pour un ratio inférieur à 1).

Les ratios d'efficacité informe sur la gestion et l'exploitation des ressources de la compagnie. Par exemple, le coefficient de rotation des stocks (coûts des marchandises vendues sur stock) indique le nombre moyen de fois que le stock est écoulé pendant l'année. Le délai moyen de rotation des stocks (365 sur coefficient de rotation des stocks) représente le nombre moyen de jours pour renouveler le stock.

Les ratios de rentabilité décrivent la profitabilité des activités de la compagnie. On y retrouve le taux de rendement des fonds propres (bénéfice net sur fonds propres), qui exprime le montant de bénéfice net généré par dollar de fonds propres. En d'autres termes, il s'agit du rendement sur l'investissement des actionnaires.

Finalement, les ratios de marché permettent de mesurer la création de valeur économique de la compagnie. Cette famille se distingue des autres en utilisant des données de marché, comme le cours boursier de la compagnie. Par exemple, le ratio valeur marchande sur valeur comptable exprime la valeur économique de la compagnie par rapport à sa valeur aux livres. Un ratio supérieur à un indique que la compagnie vaut plus maintenant que lors de l'émission d'actions.

Généralement, les analystes financiers comparent un ratio d'une compagnie à la moyenne du secteur d'activité de cette compagnie pour une année donnée. Ross et al. (2010) reportent également cette idée qui est à la base de ce mémoire. L'objectif de

notre travail est de déterminer si les ratios relatifs au secteur d'activité de la compagnie améliorent la prévision de faillite. Un ratio relatif par rapport à un benchmark bien choisi renseignerait mieux sur la difficulté/santé financière d'une compagnie qu'un ratio brut

3.3 Ratios additionnels

Nous présentons les ratios de Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984) et Chava et Jarrow (2004). Nous utilisons ces ratios additionnels dans notre investigation empirique.

Tableau 2 : Beaver (1966)		
Ratios	Interprétation	Définition
Liquidité		
Flux monétaires – dette	Proportion des flux par rapport à la dette.	Flux monétaire d'opération / Passif
Fonds de roulement – actif	Proportion du fonds de roulement par rapport à l'actif	Fonds de roulement / Actif
Fonds de roulement	Capacité d'honorer les engagements CT avec l'actif CT	Actif CT / Passif CT
Rentabilité		
Rendement de l'actif	Montant de bénéfice net par dollar d'actif	Bénéfice Net / Actif
Structure de capital		
Endettement total	Montant de dette par dollar d'actif	Passif / Actif
Efficacité		
Intervalle sans crédit	Fonds opérationnel par rapport aux dépenses opérationnelles	(Encaisse + CC – Passif CT) / (DOP – amortissement)

Dans le Tableau 2, le flux monétaire d'opération est égal au bénéfice net plus l'amortissement. Le compte client est représenté par CC. Les dépenses opérationnelles sont représentées par DOP et sont égales aux charges et dépenses administratives plus l'amortissement.

Tableau 3 : Altman (1968)		
Ratios	Interprétation	Définition
Liquidité		
Fonds de roulement – actif	Proportion du fonds de roulement par rapport à l'actif	Fonds de roulement / Actif
Bénéfice non réparti – actif	Proportion du bénéfice retenu par rapport à l'actif	Bénéfice non réparti / Actif
Rentabilité		
Rendement d'exploitation	Montant de bénéfice d'exploitation par dollar d'actif	Bénéfice avant intérêts et impôts/ Actif
Marché		
Fonds propres au marché - valeur comptable de la dette	Valeur marchande de la compagnie par rapport à la valeur aux livres de la dette	Valeur au marché des fonds propres / Passif
Efficacité		
Rotation des stocks	Nombre de fois que les stocks sont vendus	Ventes / Stock

Pour la valeur au marché des fonds propres, Altman (1968) inclut les actions ordinaires et privilégiées, alors que nous considérons seulement les actions ordinaires.

Tableau 4 : Ohlson (1980)		
Ratios	Interprétation	Définition
Structure de capital		
Taille	Mesure de la taille par l'actif	Log (Actif/Indice de prix du PNB)
Endettement total	Montant de dette par dollar d'actif	Passif / Actif
Proportion de l'actif	Indique si la dette est plus élevée que l'actif	1 si Passif > Actif; 0 sinon.
Liquidité		
Fonds de roulement – actif	Proportion du FDR par rapport à l'actif	Fonds de roulement / Actif
Fonds de roulement	Capacité d'honorer les engagements CT avec l'actif CT	Actif CT / Passif CT
Revenu d'opération – Dette	Proportion du revenu d'opération par rapport à la dette	(Revenu d'opération) / Passif
Rentabilité		
Rendement de l'actif	Montant de bénéfice net par dollar d'actif	Bénéfice Net / Actif
Tendance du bénéfice net	Indique si le bénéfice net tend à être négatif	1 si le bénéfice net est négatif deux ans de suite; 0 sinon.
Variation du bénéfice net	Mesure la variabilité/stabilité du bénéfice net	$[BN(t) - BN(t-1)] / [BN(t) - BN(t-1)]$

Dans le Tableau 4, $BN(t)$ représente le bénéfice net à l'année t . Le produit national brut est représenté par PNB. Pour la variable Taille, nous utilisons simplement le logarithme de l'actif. De plus, le revenu d'opération est estimé par le bénéfice avant intérêts, impôts et amortissement.

Tableau 5 : Zmijewski (1984)		
Ratios	Interprétation	Définition
Rentabilité		
Rendement de l'actif	Montant de bénéfice net par dollar d'actif	Bénéfice Net / Actif
Structure de capital		
Endettement total	Montant de dette par dollar d'actif	Passif / Actif
Liquidité		
Fonds de roulement	Capacité d'honorer les engagements CT avec l'actif CT	Actif CT / Passif CT

Tableau 6 : Chava et Jarrow (2004)		
Ratios	Interprétation	Définition
Rentabilité		
Rendement de l'actif	Montant de bénéfice net par dollar d'actif	Bénéfice Net / Actif
Structure de capital		
Endettement total	Montant de dette par dollar d'actif	Passif / Actif
Secteur		
Secteur	Indique à quel secteur économique la compagnie appartient	Quatre indicatrices

Dans le Tableau 6, les indicatrices pour le secteur d'activité sont établies comme suit. Les secteurs sont divisés en quatre groupes. Le premier représente le secteur de la finance, le deuxième est le secteur des transports, des communications et des utilités, le troisième contient le secteur manufacturier et le secteur minier et le quatrième contient le reste des secteurs d'activité.

Section 4 : Données

Les données comptables et financières sont issues de Compustat, gérée par *Standard & Poor's* (S&P). Compustat contient de l'information comptable et financière sur les plus grandes compagnies nord-américaines. Compustat recense presque autant de compagnies actives que de compagnies inactives depuis les années 1950. Les données sont fournies sur une base trimestrielle et annuelle et sont mises à jour mensuellement. L'information couvre le bilan, l'état des résultats et l'état des flux de trésorerie. La base contient également de l'information additionnelle de nature qualitative.

4.1 Identifiant unique

Compustat possède son propre système d'identification. La variable *Global Company Key* (GVKEY) est un identifiant unique d'une compagnie. Il reste fixe même si le nom de la compagnie ou son symbole boursier change. De plus, Compustat offre le CUSIP (*Committee on Uniform Security Identification Procedures*) comme un autre identifiant unique. Le CUSIP contient 9 caractères. Les six premiers concernent la compagnie et les deux d'après l'émission (dette ou fonds propres). Le dernier est un caractère de contrôle. Finalement, nous utilisons le couple (GVKEY, FYEAR), où FYEAR est l'année fiscale, pour identifier l'observation élémentaire (compagnie, année).

4.2 Statut de la compagnie

La variable *Research Company : Reason for Deletion* (DLRSN) renseigne sur la raison d'inactivité. DLRSN peut prendre des valeurs entre 01-20, chacune ayant une signification précise (compagnie fait fusion/acquisition, devient privée, ne remplit plus les critères du *Securities and Exchange Commission*).

La valeur 02 signifie que la compagnie a fait faillite suite à une tentative de réorganisation (*Chapter 11*), alors que la valeur 03 indique une faillite par liquidation (*Chapter 7*). DLRSN n'est pas renseignée pour les compagnies temporairement en réorganisation (sous *Chapter 11*). La variable *Assets Total – Footnotes* (AT_FN) indique alors la catégorie TL pendant la réorganisation et AG pour l'émergence (voir Tableau 7).

GVKEY	CUSIP	Nom de la compagnie	Année fiscale	DLRSN	AT_FN
062640	543881106	LORAL SPACE & COMMUNICATIONS	1996		
062640	543881106	LORAL SPACE & COMMUNICATIONS	1997		
062640	543881106	LORAL SPACE & COMMUNICATIONS	1998		
062640	543881106	LORAL SPACE & COMMUNICATIONS	1999		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2000		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2001		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2002		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2003		TL
062640	543881106	LORAL SPACE & COMMUNICATIONS	2004		TL
062640	543881106	LORAL SPACE & COMMUNICATIONS	2005		AG
062640	543881106	LORAL SPACE & COMMUNICATIONS	2006		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2007		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2008		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2009		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2010		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2011		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2012		
062640	543881106	LORAL SPACE & COMMUNICATIONS	2013		

Dans le Tableau 7, la compagnie Loral Space & Communications était en réorganisation (*Chapter 11*) en 2003 et 2004. Elle a finalement émergé en 2005 et continue d’opérer avec le même GVKEY et le même CUSIP.

4.3 Date de faillite

La variable *Research Company: Deletion Date* (DLDTE), qui indique la date d’inactivité, ne coïncide pas nécessairement avec la (vraie) date de faillite. Compustat peut reporter une date d’inactivité (ici la faillite) soit avant soit après la (vraie) date de faillite. S&P reporte la (vraie) date d’inactivité dans une base différente, Capital IQ Platform, mais HEC Montréal n’y est pas abonné.

Plusieurs bases de données renseignent sur la date de faillite. Par exemple, CRSP (*Center for Research in Security Prices*) contient les variables *Delisting Code* et *Delisting Date* et DRS (*Moody’s – Default Risk Service*) contient les variables

Bankruptcy Type et *Bankruptcy Date*. Ces bases de données présentent des problèmes similaires à Compustat.

Finalement, la base de données BankruptcyData, accessible électroniquement via son site web *www.bankruptcydata.com*, permet de compléter l'information. Cette base offre les dates de faillite des compagnies nord-américaines à partir de 1985. Le CUSIP n'est pas disponible dans cette base. On y trouve uniquement le nom de la compagnie, son secteur d'activité et sa date de faillite. Le croisement avec Compustat est alors effectué en partie manuellement. Lorsque BankruptcyData et Compustat diffèrent notablement, nous effectuons une recherche sur internet pour valider l'information.

4.4 Secteur d'activité

La variable SICH (*Standard Industrial Classification - Historical*) renseigne sur le secteur d'activité de la compagnie. Cette variable est qualitative. Elle est renseignée sur quatre positions. Chaque niveau raffine davantage la classification (voir Tableau 8 et Figure 3).

Tableau 8 : Secteurs économiques selon le <i>Standard Industrial Classification</i>	
Secteur	Description
0	Agriculture, sylviculture et pêche
1	Mines et Construction
2	Manufacture : Alimentation, tabac, textile, fourniture, produits chimiques, pétrole, vêtements, produits à base de bois, papier et imprimerie, etc.
3	Manufacture : Produits à base de métaux et plastique, machinerie industrielle, équipement de transport, équipement électronique, équipement de mesure et d'analyse, équipement médical, etc.
4	Transport, communication, gaz, électricité et services sanitaires
5	Vente au détail : Biens durables et non durables, fournitures de ménages, alimentation, etc.
6	Finance, assurance et marché immobilier
7	Services : Hôtellerie, services personnels, services pour entreprises, cinémas, divertissement, etc.
8	Service : Services juridiques, arts et musées, services de santé, éducation, recherche, gestion, etc.
9	Administration publique : justice, ordre public et sécurité, finances publiques, programmes environnementaux, programmes d'habitations, etc.

Figure 3 : Exemple de décomposition de la variable SICH

Niveau 1	1 →	Mine et construction
Niveau 2	13 →	Extraction : gaz et pétrole
Niveau 3	138 →	Services de chantiers : gaz et pétrole
Niveau 4	1382 →	Explorations de chantiers : gaz et pétrole

Le Tableau 9 illustre la différence entre la variable SIC et SICH. Alors que la variable SIC renseigne sur la classification actuelle de la compagnie (SIC = 6351), SICH renseigne sur la classification à chaque année. Par exemple, en 1987, la compagnie Acmat Corp –CL A était dans le secteur d'activité 1540 et en 1990 la compagnie était dans le secteur 6411.

GVKEY	FYEAR	CONM	SIC	SICH
001097	1987	ACMAT CORP -CL A	6351	1540
001097	1988	ACMAT CORP -CL A	6351	1540
001097	1989	ACMAT CORP -CL A	6351	1540
001097	1990	ACMAT CORP -CL A	6351	6411
001097	1991	ACMAT CORP -CL A	6351	6411
001097	1992	ACMAT CORP -CL A	6351	6411
001097	1993	ACMAT CORP -CL A	6351	6351
001097	1994	ACMAT CORP -CL A	6351	6351
001097	1997	ACMAT CORP -CL A	6351	6351
001097	1998	ACMAT CORP -CL A	6351	6351
001097	1999	ACMAT CORP -CL A	6351	6351
001097	2000	ACMAT CORP -CL A	6351	1540
001097	2001	ACMAT CORP -CL A	6351	1540
001097	2002	ACMAT CORP -CL A	6351	1540
001097	2003	ACMAT CORP -CL A	6351	1540
001097	2004	ACMAT CORP -CL A	6351	1540
001097	2005	ACMAT CORP -CL A	6351	6351
001097	2006	ACMAT CORP -CL A	6351	6351
001097	2007	ACMAT CORP -CL A	6351	6351
001097	2008	ACMAT CORP -CL A	6351	6351
001097	2009	ACMAT CORP -CL A	6351	6351
001097	2010	ACMAT CORP -CL A	6351	6351
001097	2011	ACMAT CORP -CL A	6351	6351
001097	2012	ACMAT CORP -CL A	6351	6351

Section 5 : Choix méthodologiques

5.1 Échantillonnage

Beaver (1966) effectue un appariement *un pour un déterministe* sur la base du secteur d'activité et la taille. Pour chaque compagnie en faillite, observée un certain temps avant la faillite (deux ans avant la faillite par exemple), l'auteur identifie les compagnies saines appartenant au même secteur d'activité. Parmi celles-ci, il choisit la compagnie saine avec la taille la plus proche. Beaver élimine le secteur de la finance, du transport et des utilités.

Altman (1968) effectue un appariement *un pour un déterministe randomisé* sur la base de l'année et le secteur d'activité. Pour chaque compagnie en faillite observée un certain temps avant la faillite, l'auteur identifie les compagnies saines observées à la même année et opérant dans le même secteur. Parmi celles-ci, il en sélectionne une aléatoirement. L'auteur considère seulement le secteur manufacturier.

Ohlson (1980) observe les compagnies en faillite un certain temps avant la faillite. Par la suite, pour chaque compagnie saine, il choisit au hasard une année d'observation. On peut interpréter ce mode d'échantillonnage comme un appariement *un pour plusieurs randomisé*. De plus, il retire les compagnies du secteur de la finance, du transport et des utilités.

Ce mode d'échantillonnage (l'appariement) est cohérent avec l'hypothèse d'indépendance, mais n'est pas supporté par la théorie statistique dont le mode de référence est l'échantillonnage aléatoire simple.

Shumway (2001) effectue son échantillonnage en prenant toutes les observations communes aux bases Compustat et CRSP. Il observe la faillite à l'année de faillite. En présence de données sérielles (même compagnie dans le temps), l'hypothèse d'indépendance est difficile à justifier. Shumway retire uniquement le secteur de la finance.

Tableau 10 : Modes d'échantillonnage			
Auteur	Mode	Période	Secteur
Articles			
Beaver (1966)	Appariement un pour un déterministe sur la base du secteur et la taille	1954-1964	Exclusion du secteur du transport, des utilités et du secteur de la finance
Altman (1968)	Appariement un pour un déterministe randomisé sur la base de l'année et du secteur	1946-1965	Inclusion du secteur manufacturier uniquement
Ohlson (1980)	Appariement un pour plusieurs randomisé	1970-1976	Exclusion du secteur du transport, des utilités et du secteur de la finance
Shumway (2001)	Toute la base	1962-1992	Exclusion du secteur de la finance
Notre implantation			
Beaver (1966)	Appariement un pour un déterministe sur la base du secteur et la taille	1985-2013	Exclusion du secteur du transport, des utilités et du secteur de la finance
Altman (1968)	Appariement un pour un déterministe randomisé sur la base de l'année et du secteur	1985-2013	Tous les secteurs sauf la finance
Ohlson (1980)	Appariement un pour plusieurs randomisé	1985-2013	Exclusion du secteur du transport, des utilités et du secteur de la finance
Shumway (2001)	Toute la base	1985-2013	Exclusion du secteur de la finance

5.2 Benchmark et ratios relatifs

Nous expérimentons principalement avec quatre benchmarks. Pour chaque ratio associé à un couple (compagnie, année), nous considérons la moyenne/médiane du ratio dans le secteur d'activité de cette compagnie à la même année. Par exemple, soient X un ratio financier et (ABC, 1985) le couple (compagnie, année), où ABC opère dans le secteur manufacturier. Un benchmark possible est alors la moyenne du ratio X dans le secteur manufacturier en 1985. Un autre benchmark est la médiane du ratio X dans le secteur manufacturier en 1985. Nous utilisons le secteur de niveau 1 ou 2 (voir Figure 3) pour le calcul du benchmark. Nous obtenons alors les quatre benchmarks suivants:

1. Moyenne du secteur de niveau 1 à l'année d'observation;
2. Moyenne du secteur de niveau 2 à l'année d'observation;
3. Médiane du secteur de niveau 1 à l'année d'observation;
4. Médiane du secteur de niveau 2 à l'année d'observation.

Finalement, pour le ratio brut X, le ratio relatif X_R est déterminé comme suit

$$X_R = \frac{X - X_B}{|X_B|},$$

où X_B est le benchmark choisi. Le ratio relatif devient la variable explicative dans le modèle de régression logistique et représente l'écart relatif de la compagnie au secteur d'activité à la même année d'observation.

5.3 Modèle logistique

Le modèle logistique est discuté en détail par plusieurs auteurs (Amemiya 1985, Gourieroux 1989 et Greene 2012). Soit la variable Y qui prend la valeur 1 pour une compagnie en faillite et 0 pour une compagnie saine, pour une année observée. Le modèle logistique estime la probabilité de faillite comme suit

$$P(Y = 1 | X) = F(S(X)),$$

où $X = (1, X_1, \dots, X_p)^T$ et $S(\cdot)$ est une fonction score linéaire

$$\begin{aligned} S(X) &= \beta_0 + \beta_1 X_1 + \dots + \beta_p X_p \\ &= X^T \boldsymbol{\beta}, \end{aligned}$$

où $\boldsymbol{\beta} = (\beta_0, \dots, \beta_p)^T$ est le vecteurs des coefficients de la régression logistique. Ce modèle utilise la fonction de distribution logistique

$$F(s) = \frac{e^s}{1 + e^s}.$$

Nous obtenons alors,

$$Y = \begin{cases} 1, & \text{avec probabilité } F(X^T \boldsymbol{\beta}), \\ 0, & \text{sinon} \end{cases},$$

Dans la version standard du modèle, le vecteur de coefficient $\boldsymbol{\beta}$ estimé par maximum de vraisemblance sous l'hypothèse d'indépendance. Soit Y_1, \dots, Y_n un échantillon indépendant et identiquement distribué. La fonction de vraisemblance L s'écrit

$$L(\boldsymbol{\beta} | X) = \prod_{i=1}^n [F(X_i^T \boldsymbol{\beta})]^{Y_i} [1 - F(X_i^T \boldsymbol{\beta})]^{1-Y_i}$$

et le log-vraisemblance

$$\log L(\boldsymbol{\beta} | \mathbf{X}) = \sum_{i=1}^n \left[Y_i \log F(X_i^T \boldsymbol{\beta}) + (1 - Y_i) \log (1 - F(X_i^T \boldsymbol{\beta})) \right].$$

Cette fonction est concave et admet un maximum global (McFadden 1974 et Heckman 1978). L'algorithme de Newton-Raphson permet de trouver ce maximum à l'aide des conditions de premier ordre. L'estimation par maximum de vraisemblance donne lieu à une série de tests asymptotiques de signification globale et partielle (Amemiya 1985, Gourieroux 1989 et Greene 2012). Rappelons que notre objectif est de déterminer quelles variables explicatives (ratios bruts ou ratios relatifs) offrent une meilleure prévision et non de juger entre l'efficacité de différents modèles économétriques. C'est pourquoi nous travaillons uniquement avec le modèle logistique, mais avec des variables explicatives différentes.

5.4 Indicateurs de performance

Suite à l'estimation du modèle, nous posons

$$\hat{Y}_i = \begin{cases} 1, & \text{si } \hat{P}(Y_i = 1 | X_i) > p^* \\ 0, & \text{sinon} \end{cases},$$

où $\hat{P}(Y_i = 1 | X_i) = F(X_i^T \hat{\boldsymbol{\beta}})$ et p^* est un seuil de coupure donné. Souvent, les auteurs déterminent le seuil p^* qui maximise le nombre d'individus bien classés. Nous considérons les deux indicateurs de performance suivants : le coût moyen d'une mauvaise affectation (CMMA) et le ROC (*Receiver Operating Characteristic*). Ces indicateurs sont déterminés soit en-échantillon (*in-sample*) soit hors-échantillon (*out-of-sample*). Alors que les exemples suivants sont basés sur une prévision en-échantillon, notre investigation empirique utilise une prévision hors-échantillon.

Le coût moyen d'une mauvaise affectation (CMMA) permet de tenir compte de l'impact économique d'une erreur de classification. Par exemple, pour un seuil de coupure donné ($p^* = 0.5$), nous obtenons le tableau croisé $Y \times \hat{Y}$ suivant :

Tableau 11 : Y x \hat{Y}				
		\hat{Y}		Total
		1	0	
Y	1	124	72	196
	0	97	99	196
Total		221	171	392

À l'aide de ce tableau, le CMMA s'écrit

$$\begin{aligned}
 \text{CMMA} &= c \times \hat{P}(Y = 1 \text{ et } \hat{Y} = 0) + 1 \times \hat{P}(Y = 0 \text{ et } \hat{Y} = 1) \\
 &\quad + 0 \times \hat{P}(Y = 1 \text{ et } \hat{Y} = 1) + 0 \times \hat{P}(Y = 0 \text{ et } \hat{Y} = 0) \\
 &= c \times \hat{P}(Y = 1 \text{ et } \hat{Y} = 0) + \hat{P}(Y = 0 \text{ et } \hat{Y} = 1),
 \end{aligned}$$

où c représente le coût d'une mauvaise affectation d'une compagnie en faillite (erreur de Type I), sous l'hypothèse que le coût d'une mauvaise affectation d'une compagnie saine (erreur de Type II) est unitaire. Le multiplicateur c reflète le coût d'une erreur de Type I par rapport à l'erreur de Type II. Nous expérimentons avec $c \in \{1; 2; 3; 4; 5\}$. Par exemple, pour un seuil $p^* = 0.5$ et $c = 2$, le Tableau 11 indique

$$\text{CMMA} = 2 \times \frac{72}{392} + \frac{97}{392} = 0.6148.$$

Le ROC, développé durant la Deuxième Guerre mondiale pour juger de l'efficacité des radars à détecter des objets ennemis, est encore fréquemment utilisé pour évaluer la performance d'un modèle binaire. Dans notre contexte, le ROC repose sur l'idée simple qu'un modèle acceptable a une meilleure capacité à prévoir la faillite parmi les compagnies en faillite que parmi les compagnies saines et, ce, pour chaque seuil de coupure p^* .

Nous définissons la sensibilité comme la capacité du modèle à prévoir la faillite parmi les compagnies en faillite et la spécificité comme la capacité du modèle à prévoir la survie parmi les compagnies saines. Donc, $1 - \text{spécificité}$ est la (mauvaise) capacité du modèle à prévoir la faillite parmi les compagnies saines. Par exemple, pour un seuil de coupure donnée ($p^* = 0.5$), le Tableau 11 indique

$$\begin{aligned}\text{Sensibilité} &= P(\hat{Y} = 1 \mid Y = 1) \\ &= \frac{124}{124 + 72} = 63.27\%\end{aligned}$$

$$\begin{aligned}\text{Spécificité} &= P(\hat{Y} = 0 \mid Y = 0) \\ &= \frac{99}{97 + 99} = 50.51\%\end{aligned}$$

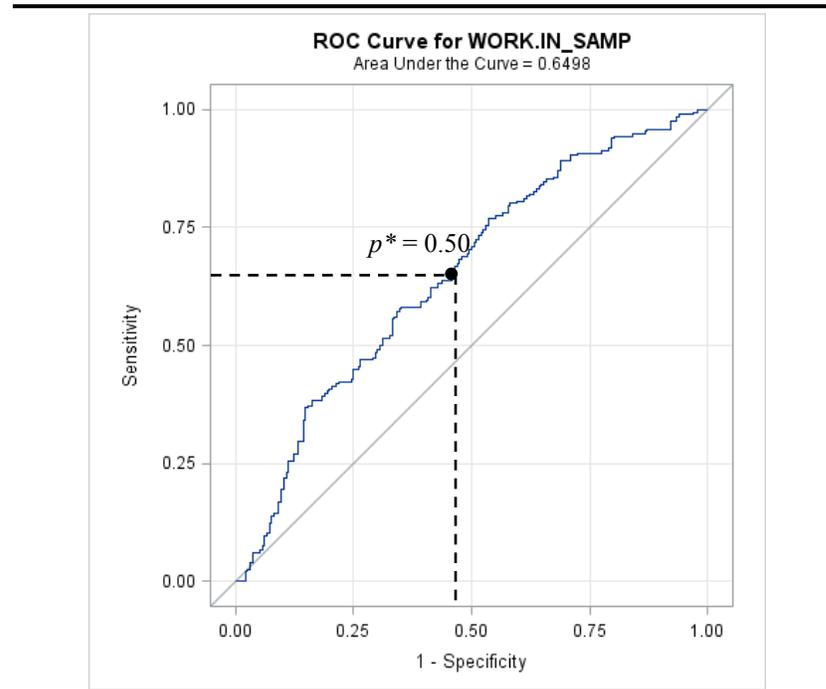
$$\begin{aligned}1 - \text{Spécificité} &= P(\hat{Y} = 1 \mid Y = 0) \\ &= \frac{97}{97 + 99} = 49.49\%.\end{aligned}$$

En toute logique, un modèle acceptable vérifie

$$\text{Sensibilité} > 1 - \text{Spécificité}$$

La Figure 4 illustre le point sur la courbe ROC associé à cet exemple.

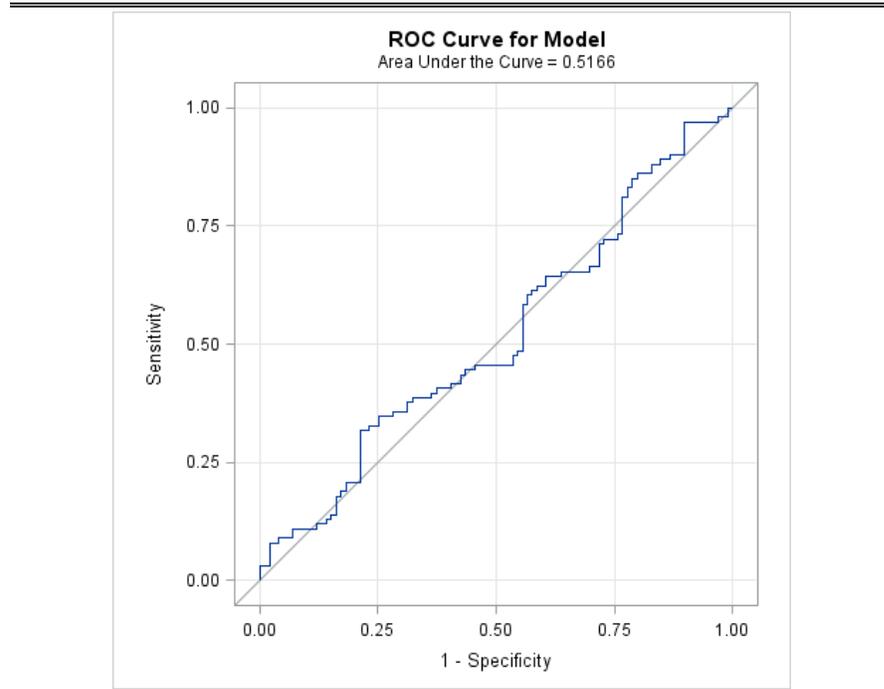
Figure 4 : Point sur la courbe ROC au seuil $p^* = 0.50$



En faisant varier $p^* \in [0, 1]$, nous obtenons la courbe ROC. Plus elle est au-dessus de la diagonale, meilleur est le modèle. SAS indique également l'aire sous la courbe ROC. La diagonale reflète un modèle sans pouvoir prévisionnel. La Figure 5 présente la courbe

ROC pour un modèle où la variable explicative X est indépendante de la variable Y. L'Annexe 2 présente la sensibilité et la spécificité pour plusieurs seuils de coupure, telles que reportées par SAS.

Figure 5 : Courbe ROC – X et Y indépendants



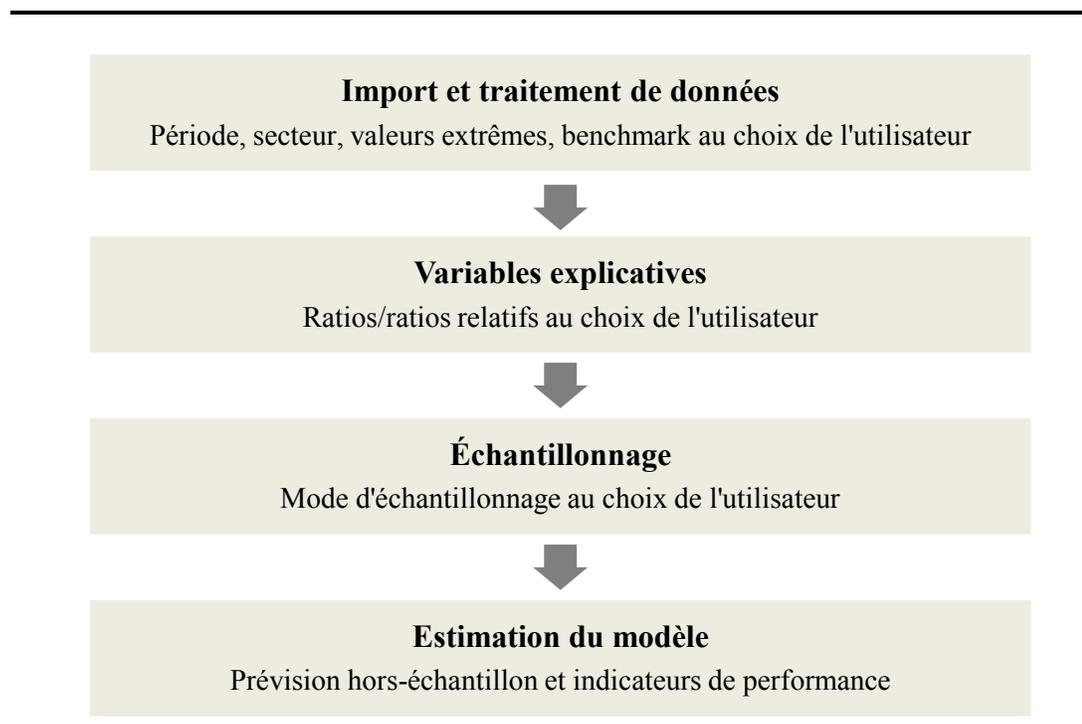
Section 6 : Investigation empirique

Notre programme informatique est développé sous SAS. L'éditeur SAS permet de lire des fichiers de taille virtuellement infinie. SAS SQL permet une gestion efficace de fichiers de données. SAS STAT permet d'utiliser de tous les modèles statistiques reconnus. De plus, SAS est doté d'un langage de programmation qui permet d'automatiser toutes les tâches à partir du traitement de données, à l'analyse statistique jusqu'à la gestion des sorties.

Notre programme contient près de 2000 lignes de code. Il a nécessité un effort étalé sur une période de six mois. Il permet de gérer la base Compustat : importer, transformer, purifier les données et extraire des échantillons de compagnies en faillite et saines observées à travers une batterie de ratios selon un mode d'échantillonnage présélectionné (voir la section 5.1) et une sélection de ratios présélectionnée (voir la section 3.3). Le programme permet aussi de coupler Compustat avec la base BankruptcyData. En plus, il permet de transformer les ratios bruts en ratios relatifs où chaque compagnie est comparée à son secteur d'activité. Le benchmark est laissé au choix de l'utilisateur. Le programme permet finalement de rouler des paires de modèles avec ratios bruts et ratios relatifs et de les comparer sur la base du ROC et du CMMA (voir la section 5.4). En tout, le programme permet de comparer près de 640 couples de modèles issues de cinq sélections de ratios, quatre modes d'échantillonnage, deux traitements de valeurs extrêmes, quatre choix de benchmark et quatre choix de période (toute la période 1985-2013, la période sans les crises financières, seulement les crises financières et le choix de fenêtres de temps). De plus, ce programme permet d'effectuer des analyses sectorielles. Nous avons mené une analyse de plus d'une dizaine de secteurs économiques. Les modèles sont comparés sur la base de six critères (ROC et CMMA avec $c = \{1, 2, 3, 4, 5\}$).

La Figure 6 illustre les grandes sections de notre programme.

Figure 6 : Sections de notre programme informatique



6.1 Paramètres de flexibilité

L'analyse est effectuée selon les paramètres de flexibilité choisis par l'utilisateur. Ce dernier peut choisir la période d'étude entre 1985 et 2013. Par exemple, il peut éliminer les crises financières durant la période. Ensuite, l'analyste choisit l'horizon de prévision. Il spécifie le décalage entre la date de faillite et la date d'observation. Par exemple, pour un horizon de prévision de deux ans, l'observation des compagnies en faillite se fait deux ans avant la faillite. L'utilisateur a aussi le choix du secteur d'activité, de la proportion de valeurs extrêmes à éliminer et du benchmark pour les ratios relatifs.

Finalement, il peut choisir les ratios et modes d'échantillonnage des auteurs suivants : Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984), Shumway (2001), Chava et Jarrow (2004) et Ross et al. (2010). Le croisement de ces options aboutit à des centaines de modèles. La section qui suit présente les résultats d'un modèle arbitraire.

Nous partageons l'échantillon en deux parties égales. Alors que la première sert à l'estimation du modèle, la deuxième sert à la prévision hors-échantillon.

6.2 Résultats pour un modèle arbitraire

En guise d'exemple, nous utilisons le mode d'échantillonnage d'Altman (1968), les ratios de Zmijewski (1984), la période 1985-2013 et un horizon de prévision de deux ans. Nous excluons le secteur de la finance et nous éliminons 1% des valeurs extrêmes (sur chaque aile). Le benchmark est basé sur la moyenne et le secteur d'activité de niveau 1. L'échantillon final contient 784 compagnies dont 392 faillites et 392 compagnies saines appariées sur la base de l'année et du secteur. La partie hors-échantillon contient 392 compagnies dont 196 faillites et 196 compagnies saines appariées.

Pour des fins de présentation, les ratios brut/relatifs de Zmijewski (1984) sont renommés de manière à reconnaître leur famille (voir Tableau 12).

Tableau 12 : Zmijewski (1984)		
Abréviations	Ratios	Famille
R1	Bénéfice Net / Actif	Rentabilité
D1	Passif / Actif	Structure de capital
L1	Actif CT / Passif CT	Liquidité

Le Tableau 13 et le Tableau 14 contiennent des statistiques descriptives pour chacun des ratios par groupe de compagnies (faillite/saines). Le Tableau 15 et le Tableau 16 présentent les résultats du modèle logistique.

Tableau 13 : Statistiques descriptives						
Ratios bruts	2 ans avant la faillite			Compagnies saines		
	N	Moyenne	Écart-type	N	Moyenne	Écart-type
R1	392	-0.31	0.56	392	-0.16	0.58
D1	392	0.78	0.54	392	0.61	0.57
L1	392	1.85	1.75	392	2.59	3.05
Ratios relatifs	2 ans avant la faillite			Compagnies saines		
	N	Moyenne	Écart-type	N	Moyenne	Écart-type
R1	392	-2.47	6.99	392	0.44	5.79
D1	392	0.34	0.92	392	0.06	1.00
L1	392	-0.25	0.72	392	0.04	1.19

Tableau 14 : Quartiles des distributions						
Ratios bruts	2 ans avant la faillite			Compagnies saines		
	Q1	Q2	Q3	Q1	Q2	Q3
R1	-0.36	-0.12	-0.01	-0.11	0.02	0.07
D1	0.51	0.73	0.91	0.35	0.53	0.71
L1	0.92	1.42	2.17	1.05	1.72	2.74
Ratios relatifs	2 ans avant la faillite			Compagnies saines		
	Q1	Q2	Q3	Q1	Q2	Q3
R1	-3.57	-0.51	0.9	-0.06	1.25	2.08
D1	-0.12	0.22	0.58	-0.39	-0.09	0.19
L1	-0.62	-0.41	-0.15	-0.55	-0.3	0.13

Les ratios bruts indiquent que les compagnies en faillite sont en moyenne moins rentables (R1), plus endettées (D1) et dispose de moins de liquidité (L1) que les compagnies saines. Les ratios relatifs présentent les mêmes caractéristiques, lesquels sont cohérents avec la théorie financière.

Tableau 15 : Tests de signification globale				
Tests	Ratios bruts		Ratios relatifs	
	Valeur	Valeur- <i>p</i>	Valeur	Valeur- <i>p</i>
Ratio de vraisemblance	15.55	0.0014	28.79	<.0001
Score khi-2 (Lagrange)	14.60	0.0022	24.53	<.0001
Wald khi-2	13.06	0.0045	20.76	0.0001

Tableau 16 : Estimation du modèle								
Ratios	Ratios bruts				Ratios relatifs			
	Estimation	Écart-type	Wald khi-2	Valeur- <i>p</i>	Estimation	Écart-type	Wald khi-2	Valeur- <i>p</i>
Constante	0.1196	0.2096	0.3255	0.5683	-0.1113	0.1076	1.0700	0.3009
R1	-0.3059	0.1838	2.7691	0.0961	-0.0826	0.0236	12.2777	0.0005
D1	0.0772	0.1829	0.1782	0.6730	0.00566	0.1091	0.0027	0.9586
L1	-0.1116	0.0451	6.1335	0.0133	-0.2578	0.1202	4.5981	0.0320

Les trois tests de signification globale sont asymptotiquement équivalents, mais peuvent présenter des conclusions différentes pour un échantillon de taille faible. Une valeur-*p* inférieure au seuil de signification $\alpha = 5\%$ indique le rejet de l'hypothèse nulle

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

et donc la signification du modèle. Les trois tests du Tableau 14 indiquent que le modèle est globalement significatif avec les ratios bruts et avec les ratios relatifs. Le modèle avec ratios relatifs est d'ailleurs plus significatif (valeurs- p plus faibles pour les trois tests) que le modèle avec ratios bruts.

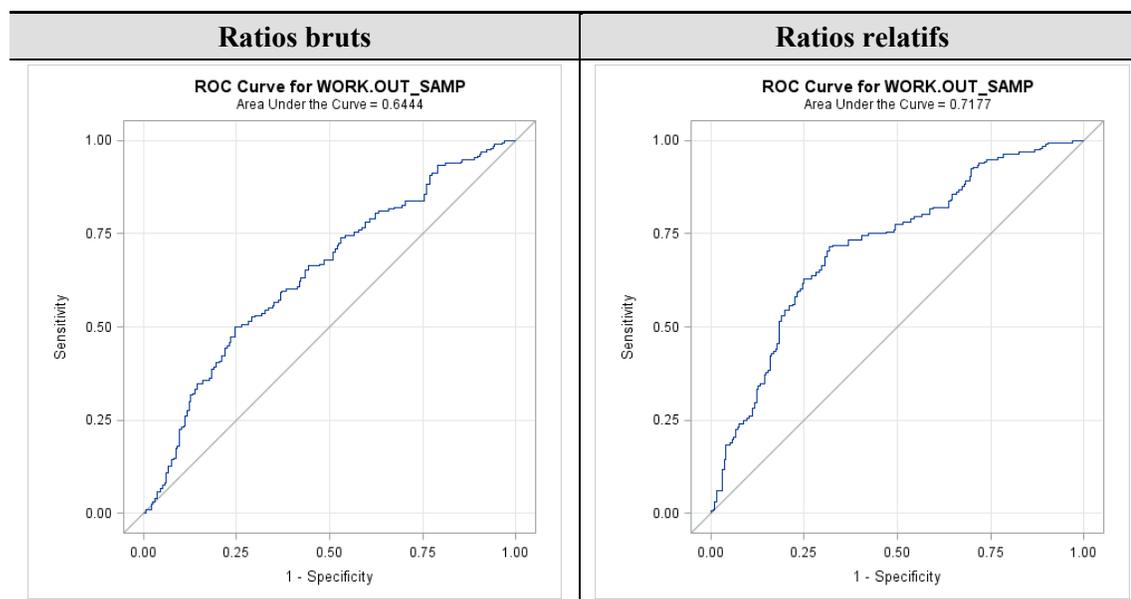
De plus, en utilisant les ratios bruts, seulement L1 est significatif, alors que pour les ratios relatifs, R1 et L1 sont significatifs (au seuil de signification de 5%). Les signes des estimations sont cohérents avec la théorie financière. L'estimation de la variable R1 (ratios relatifs) est de -0.0826. Une augmentation de la rentabilité diminue alors la probabilité de faillite, toute chose étant égale par ailleurs. L'estimation de la variable L1 (ratios relatifs) est de -0.2578. Une augmentation de la liquidité diminue aussi la probabilité de faillite, toute chose étant égale par ailleurs. La variable D1 n'est pas significative. Cela peut s'expliquer par le fait que les valeurs comptables ne reflètent pas nécessairement les valeurs économiques et D1 ne permet pas de discriminer entre compagnie en faillite et compagnie saine. Cependant, le signe de l'estimation reste cohérent avec la théorie financière.

Pour le modèle avec les ratios bruts, le CMMA minimum est de 47.45% et est atteint au seuil de coupure de 0.45, alors que le modèle avec les ratios relatifs offre un CMMA minimum de 44.64% au seuil de 0.40 (voir Tableau 17). Ce qui représente une légère amélioration (coût plus faible) avec l'usage des ratios relatifs.

Tableau 17 : CMMA (en %) par seuil de coupure ($c = 2$)		
Seuil de coupure	Ratios bruts	Ratios relatifs
0.20	49.74	49.23
0.25	49.23	48.47
0.30	48.47	48.21
0.35	48.72	47.19
0.40	49.23	44.64
0.45	47.45	47.70
0.50	57.14	55.61
0.55	89.03	77.30
0.60	98.21	83.93
0.65	99.23	91.33
0.70	99.23	95.15
0.75	99.23	97.19
0.80	99.23	97.96

La Figure 7 présente les courbes ROC pour le modèle avec ratios bruts et le modèle avec ratios relatifs. La courbe du modèle avec ratios relatifs est plus élevée et l'aire sous la courbe est de 0.7177 comparée à 0.6444 pour le modèle avec les ratios bruts. Cela indique que le modèle avec ratios relatifs performe mieux.

Figure 7 : Courbes ROC



Finalement, dans ce modèle, l'usage des ratios relatifs améliore la qualité de la prévision, mesurée par le CMMA et le ROC. Malheureusement, ce résultat n'est pas robuste. Les résultats sont mitigés. Les ratios relatifs performant parfois mieux, mais sont parfois dominés par les ratios bruts.

6.3 Résultats globaux

Notre investigation empirique est basée sur les 640 couples de modèles décrits à la section 6. Chaque modèle est roulé avec les ratios bruts et avec les ratios relatifs. Le couple de modèles est comparé sur la base de six critères, le ROC et le CMMA avec $c = \{1, 2, 3, 4, 5\}$. Les résultats sont mitigés. Les ratios bruts performant légèrement mieux que les ratios relatifs. Néanmoins, l'avantage pour les ratios bruts est en général faible au cas par cas.

Section 7 : Conclusion

Nous investiguons l'idée que les ratios relatifs au secteur d'activité de la compagnie permettent une meilleure prévision de faillite que les ratios bruts.

Nous considérons les ratios, les modes d'échantillonnage, le traitement des valeurs extrêmes et les indicateurs de performance proposés par Beaver (1966), Altman (1968), Ohlson (1980), Zmijewski (1984), Shumway (2001), Chava et Jarrow (2004) ainsi que Ross et al. (2010). Nous proposons également plusieurs choix pour le benchmark. Nous offrons une grande flexibilité dans le choix du modèle.

Nous mesurons la performance des modèles à l'aide du CMMA et du ROC. Nous comparons le modèle avec ratios bruts et avec ratios relatifs. Notre analyse aboutit à des résultats mitigés. Les ratios relatifs assurent parfois une meilleure prévision, mais sont parfois dominés par les ratios bruts. Malgré ces résultats, nous croyons encore que les ratios relatifs sont meilleurs. Nous arrivons à la conclusion que la difficulté réside dans le choix du benchmark. Ce dernier est construit sur la base du secteur économique, qui reflète l'activité principale de la compagnie. Comme nous opérons à même la base Compustat, avec les plus grandes compagnies boursières en Amérique du Nord, un benchmark basé sur l'activité principale n'est pas nécessairement représentatif des activités multiples de la compagnie.

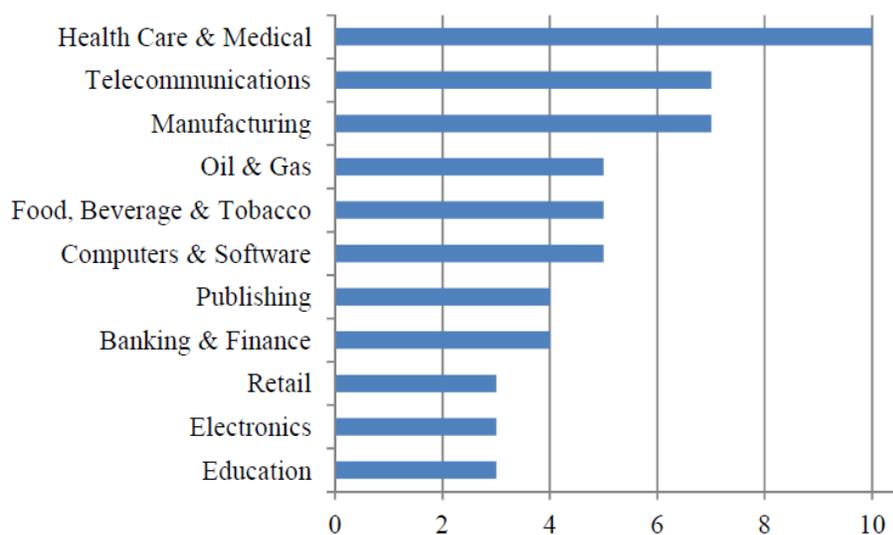
Annexe 1 : Données sur la faillite corporative

Le Tableau 18 renseigne sur le nombre de faillites et la valeur des actifs des plus grandes compagnies publiques américaines sur la période 2010-2013.

Année	Nombre de faillites	Actif Total (millions \$)	Actif moyen (millions \$)
2010	106	89,109	840
2011	86	103,990	1,209
2012	87	70,843	824
2013	71	42,626	600

La Figure 8 indique le nombre de faillites par secteur d'activité en 2013.

Figure 8 : Nombre de faillites par secteur d'activité (2013)



Source: <http://www.turnaroundletter.com/blogs/15-bankruptcy-news/post/28801-bankruptcy-news-chapter-11-recap-forecast-prepackaged-bankruptcy-filings-rise>

Dans le Tableau 19, on retrouve la liste des 10 plus grandes faillites de 2013. Leur actif combiné représente près de 30 milliards de dollars.

Tableau 19 : Liste des 10 plus grandes faillites de 2013		
Compagnie	Date de faillite	Actif (millions \$)
Cengage Learning, Inc.	02-07-2013	7,505
Penson Worldwide, Inc.	11-01-2013	6,197
Dex One Corporation	17-03-2013	2,835
Excel Maritime Carriers Ltd.	01-07-2013	2,722
Anchor Bancorp Wisconsin Inc.	12-08-2013	2,368
Exide Technologies	10-06-2013	2,195
Central European Distribution Corporation	07-04-2013	2,073
RDA Holding Co. (Reader's Digest)	17-02-2013	1,564
SuperMedia, Inc.	18-03-2013	1,410
Revel AC, Inc.	25-08-2013	1,170

Annexe 2 : Sensibilité et spécificité

Le Tableau 20 représente une sortie SAS illustrant le calcul de la sensibilité et de la spécificité. Notons que cette sortie est disponible pour la portion de l'échantillon servant à l'estimation du modèle (*in-sample*) seulement.

Tableau 20 : Sensibilité et spécificité									
Seuil de coupure	Correct		Incorrect		Pourcentages				
	Év.	Non-Év	Év.	Non-Év	Correct	Sens.	Spéc.	Faux positif	Faux négatif
0.20	195	5	191	1	51.0	99.5	2.6	49.5	16.7
0.25	194	6	190	2	51.0	99.0	3.1	49.5	25.0
0.30	190	12	184	6	51.5	96.9	6.1	49.2	33.3
0.35	188	16	180	8	52.0	95.9	8.2	48.9	33.3
0.40	186	27	169	10	54.3	94.9	13.8	47.6	27.0
0.45	178	45	151	18	56.9	90.8	23.0	45.9	28.6
0.50	124	99	97	72	56.9	63.3	50.5	43.9	42.1
0.55	42	173	23	154	54.8	21.4	88.3	35.4	47.1
0.60	21	182	14	175	51.8	10.7	92.9	40.0	49.0
0.65	11	185	11	185	50.0	5.6	94.4	50.0	50.0
0.70	5	189	7	191	49.5	2.6	96.4	58.3	50.3
0.75	1	190	6	195	48.7	0.5	96.9	85.7	50.6
0.80	0	191	5	196	48.7	0.0	97.4	100.0	50.6

Dans le Tableau 20, la colonne Correct indique une bonne affectation ($Y = \hat{Y}$). À l'inverse, la colonne Incorrect indique une mauvais affectation ($Y \neq \hat{Y}$). Un Événement (Év.) indique une faillite, alors qu'un Non-événement (Non-Év.) indique une compagnie saine. Par exemple, au seuil de coupure de 0.5, les pourcentages sont calculés comme suit :

$$\begin{aligned} \text{Correct} &= \hat{P}(Y = 1 \text{ et } \hat{Y} = 1) + \hat{P}(Y = 0 \text{ et } \hat{Y} = 0) \\ &= \frac{124 + 99}{392} = 56.89\% \end{aligned}$$

$$\begin{aligned} \text{Sensibilité} &= \hat{P}(\hat{Y} = 1 \mid Y = 1) \\ &= \frac{124}{124 + 72} = 63.27\% \end{aligned}$$

$$\begin{aligned}\text{Spécificité} &= \hat{P}(\hat{Y} = 0 \mid Y = 0) \\ &= \frac{99}{97 + 99} = 50.51\%\end{aligned}$$

$$\begin{aligned}\text{Faux positifs} &= \hat{P}(Y = 0 \mid \hat{Y} = 1) \\ &= \frac{97}{124 + 97} = 43.89\%\end{aligned}$$

$$\begin{aligned}\text{Faux négatifs} &= \hat{P}(Y = 1 \mid \hat{Y} = 0) \\ &= \frac{72}{72 + 99} = 42.11\%\end{aligned}$$

Bibliographie

1. Altman, E.I. (1968). Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance* (23): 589-609.
2. Altman, E.I., Haldeman, R.G. et Narayanan, P. (1977). ZETA Analysis: A New Model to Identify Bankruptcy Risk of Corporations. *Journal of Banking and Finance* (1): 29-54.
3. Altman, E.I., Loris, B. (1976). A Financial Early Warning System for Over-the-Counter Broker-Dealers. *Journal of Finance* (31): 1201-1217.
4. Altman, E.I., Marco, G. et Varetto, F. (1994). Corporate Distress Diagnosis: Comparison Using Linear Discriminant Analysis and Neural Networks (The Italian Experience). *Journal of Banking and Finance* (18): 505-529.
5. Altman, E.I. et McGough, T.P. (1974). Evaluation of a Company as a Going Concern. *Journal of Accountancy* (138): 50-57.
6. Altman, E.I. et Narayanan, P. (1997). An International Survey of Business Failure Classification Models. *Financial Markets, Institutions & Instruments* (6): 1-57.
7. Amemiya, T. (1985), *Advanced Econometrics*, Harvard University Press, Cambridge.
8. BankruptcyData (2013). 2013 Public Company Bankruptcy Filings Annual Report. <http://www.bankruptcydata.com/Research/ProprietaryResearch.htm>
9. Barros, C.P., Ferreira, C. et Williams, J. (2007). Analysing the Determinants of Performance of Best and Worst European Banks: A Mixed Logit Approach. *Journal of Banking and Finance* (31): 2189-2203.
10. Beaver, W.H. (1966). Financial Ratios as Predictors of Failure. *Journal of Accounting Research* (4): 71-111.
11. Bishop, C.M. (1994). Neural Networks and their Applications. *Review of Scientific Instrument* (65): 1803-1833.
12. Blum, M. (1974). Failing Company Discriminant Analysis. *Journal of Accounting Research* (12): 1-25.
13. Chava, S. et Jarrow, R.A. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance* (8): 537-569.

14. Cheng, K.F., Chu, C.K. et Hwang, R. (2010). Predicting Bankruptcy Using Discrete-Time Semiparametric Hazard Model. *Quantitative Finance* (10): 1055-1066.
15. Deakin, E.B. (1972). A Discriminant Analysis of Predictors of Business Failure. *Journal of Accounting Research* (10): 167-179.
16. Du Jardin, P. (2010). Predicting Bankruptcy Using Neural Networks and Other Classification Methods: The Influence of Variable Selection Techniques on Model Accuracy. *Neurocomputing* (73): 2047-2060
17. Edmister, R.O. (1972). An Empirical Test of Financial Ratio Analysis for Small Business Failure Prediction. *Journal of Financial and Quantitative Analysis* (7): 1477-1493.
18. Gourieroux, C. (1989). *Économétrie des variables qualitatives*, Economica, 2^e édition, Paris.
19. Greene, W.H. (2012). *Econometric Analysis*. Prentice Hall, 7^e edition, Boston.
20. Grice, J.S. et Ingram, R.W. (2001). Tests of Generalizability of Altman's Prediction Model. *Journal of Business Research* (54): 53-61.
21. Heckman, J.J. (1978). Dummy Endogenous Variables in a Simultaneous Equation System. *Econometrica* (46): 931-959.
22. Hensher, D.A. et Jones, S. (2007). Forecasting Corporate Bankruptcy: Optimizing the Performance of the Mixed Logit Model. *Abacus* (43): 240-264.
23. Hensher, D.A., Jones, S. et Greene, W.H. (2007). An Error Component Logit Analysis of Corporate Bankruptcy and Insolvency Risk in Australia. *The Economic Record* (83): 86-103.
24. Lang, L.H.P. et Stulz R. (1992). Contagion and Competitive Intra-Industry Effects of Bankruptcy Announcements. *Journal of Financial Economics* (32): 45-60.
25. Libby, R. (1975). Accounting Ratios and the Prediction of Failure: Some Behavioral Evidence. *Journal of Accounting Research* (13): 150-161.
26. Maksimovic, V. et Phillips, G. (1998). Asset Efficiency and Reallocation Decisions of Bankrupt Firms. *Journal of Finance* (52): 1495-1532.
27. McFadden, D. (1974). Conditional Logit Analysis of Qualitative Choice Behavior, édité par Zarembka, P. (1974). *Frontiers in Econometrics*, Academic Press, New York, 105-142.

28. Merton, R.C. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance* (29): 449-470.
29. Moyer, R.C. (1977). Forecasting Financial Failure: A Re-Examination. *Financial Management* (6): 11-17.
30. Nam, C.W., Kim, T.S, Park, N.J. et Lee, H.K. (2008). Bankruptcy Prediction Using a Discrete-Time Duration Model Incorporating Temporal and Macroeconomic Dependencies. *Journal of Forecasting* (27): 493-506.
31. Ohlson, J.A. (1980). Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research* (18): 109-131.
32. Opler, T.C. et Titman, S. (1994). Financial Distress and Corporate Performance. *Journal of Finance* (49): 1015-1040.
33. Ross, S.A. et al. (2010). *Fundamentals of Corporate Finance*, McGraw-Hill Ryerson, 7th edition, Ontario.
34. Shleifer A. et Vishny, R.W. (1992). Liquidation Values and Debt Capacity : A Market Equilibrium Approach. *Journal of Finance* (47): 1343-1366.
35. Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business* (74): 101-124.
36. Sinha, A.P. et May, J.H. (2004). Evaluating and Tuning Predictive Data Mining Models Using Receiver Operating Characteristic Curves. *Journal of Management Information Systems* (21): 249-280.
37. Stein, R.M. (2005). The Relationship Between Default Prediction and Lending Profits: Integrating ROC Analysis and Loan Pricing. *Journal of Banking and Finance* (29): 1213-1236.
38. Stein, R.M. (2007). Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation. *Journal of Risk Model Validation* (1): 77-113.
39. Tsai, B. (2013). An Early Warning System of Financial Distress Using Multinomial Logit Models and a Bootstrapping Approach. *Emerging Markets Finance & Trade* (49): 43-69.
40. Wilcox, J.W. (1973). A Prediction of Business Failure Using Accounting Data. *Journal of Accounting Research* (11): 163-179.
41. Wu, Y., Gaunt, C. et Gray, S. (2010). A Comparison of Alternative Bankruptcy Models. *Journal of Contemporary Accounting & Economics* (6): 34-45.

42. Zhou, L. (2013). Performance of Corporate Bankruptcy Prediction Models on Imbalanced Dataset: The Effect of Sampling Methods. *Knowledge-Based Systems* (41): 16-25.
43. Zmijewski, M.E. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models. *Journal of Accounting Research* (22): 59-82.